MATH 106 Workbook
Version 1.1

August 27, 2021

Contents

About This Workbook

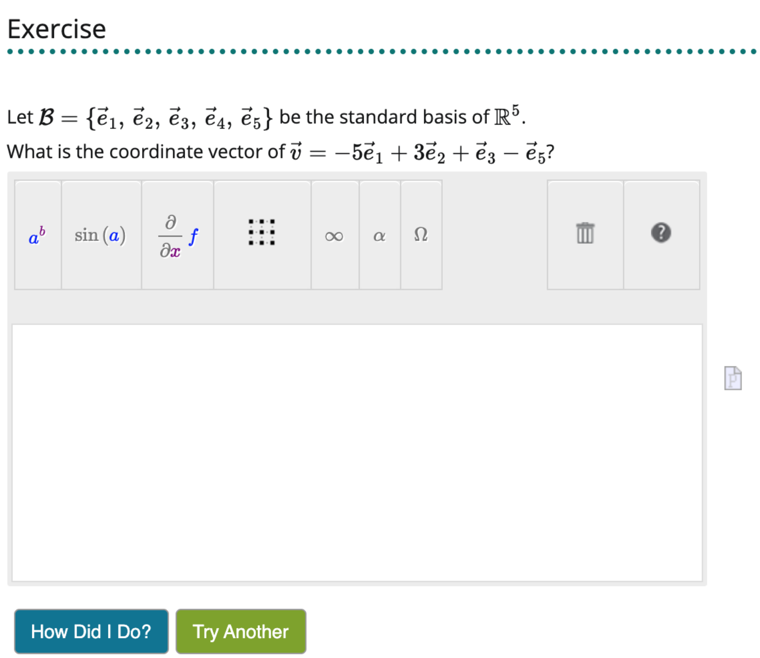
This workbook is intended as a companion to the Mobius courseware for MATH 106. This workbook is not intended to replace the online courseware. It is expected that you are working through the material online and taking notes in the workbook as you go.

The structure of this workbook, into units and lessons, follows the structure in Mobius. One exception is that the workbook does not include Unit 0. We recommend you complete Unit 0 in Mobius. You can easily navigate to units and lessons in this workbook by clicking the appropriate link in the workbook's Table of Contents.

The courseware in Mobius contains several interactive elements. Here is how they will appear in your workbook:

Exercises

In Mobius, Exercise questions are opportunities for you to practice what you have just learned. These questions come in many formats including multiple choice and numerical entry.



An Exercise question in Mobius

In this workbook, Exercise questions will appear as rounded boxes with black borders and the heading *"A question appears in Mobius"*:

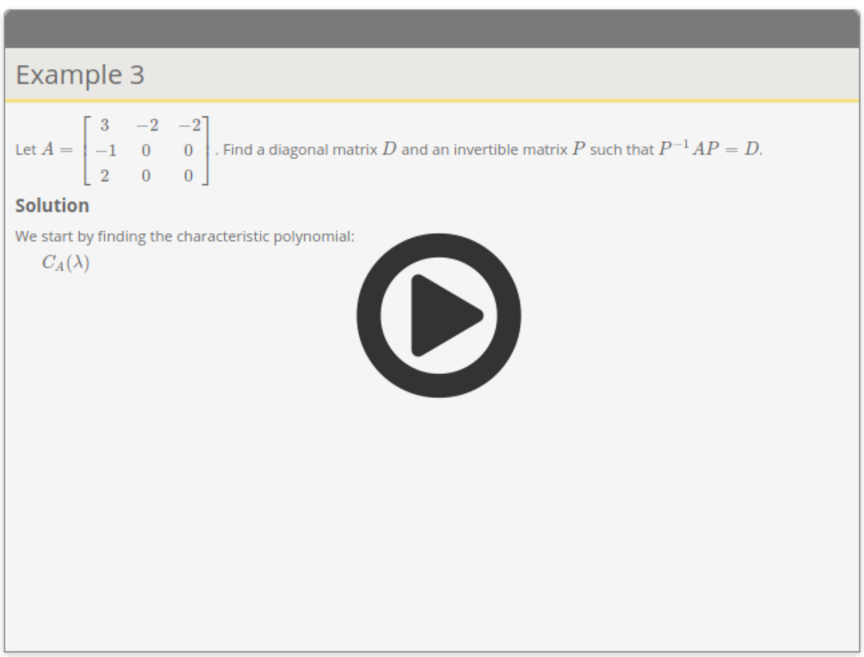
A question appears in Mobius

An Exercise question in the workbook

You can use the space included to write down the corresponding question from Mobius and your solution. We recommend that you practice entering your final answer into Mobius so that you familiarize yourself with Mobius syntax.

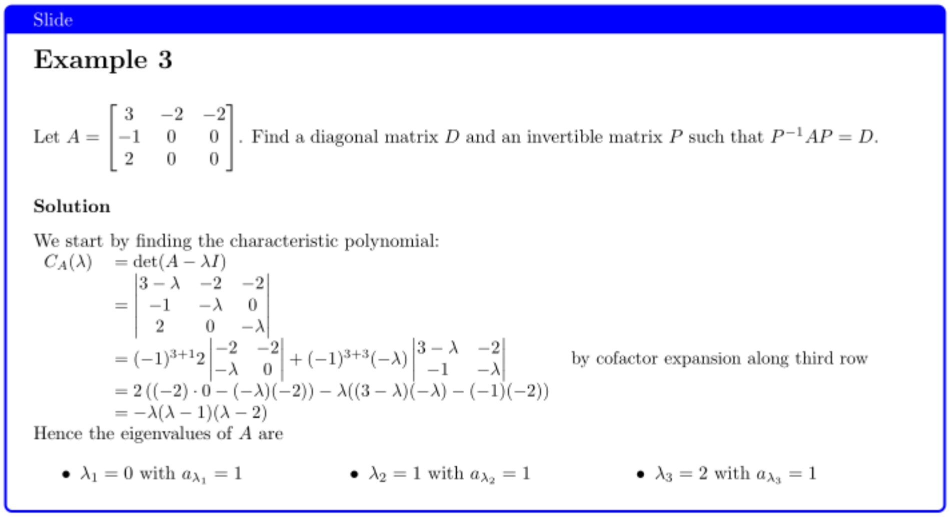
Slideshows

In Mobius, slideshows are typically narrated videos that review examples or key concepts.



A slide in Mobius

In your workbook, each slide of a slideshow will appear inside a rounded box with blue borders. The beginning of the slideshow is indicated by the words *"A slideshow appears in Mobius"*.

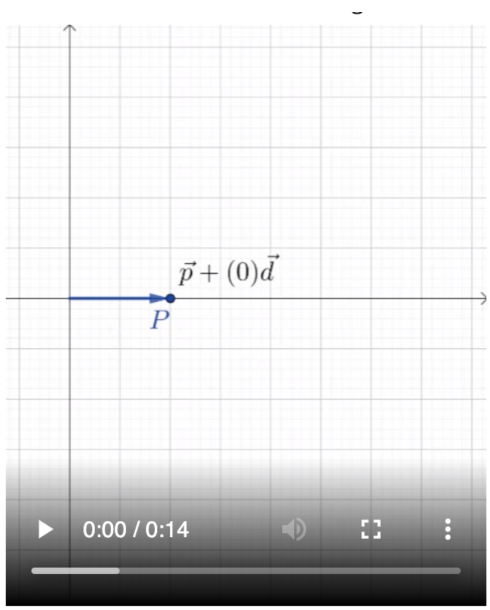


A slide in the workbook

Your workbook shows the final build of each slide. We recommend that you play the slideshow in Mobius and follow along, adding notes on the slides in your workbook.

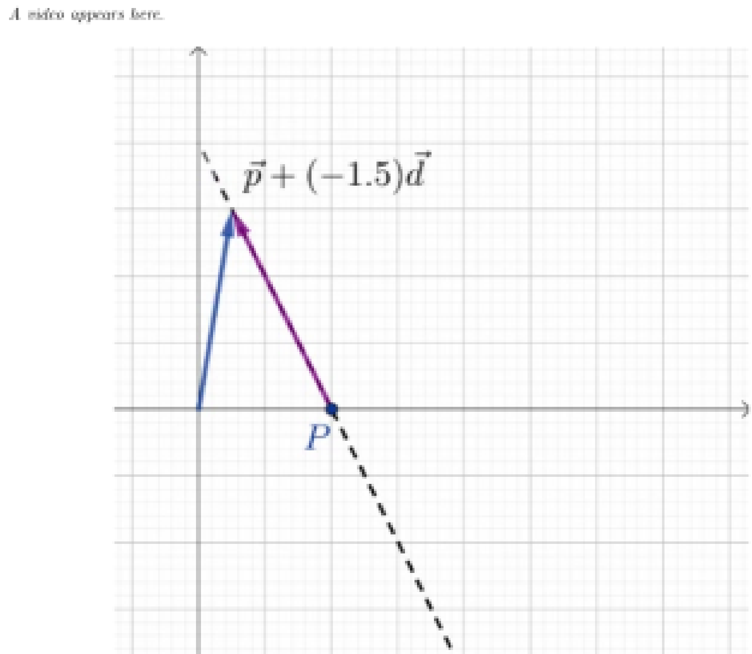
Videos

In Mobius, videos may appear within slideshows and as standalone objects (usually without audio).



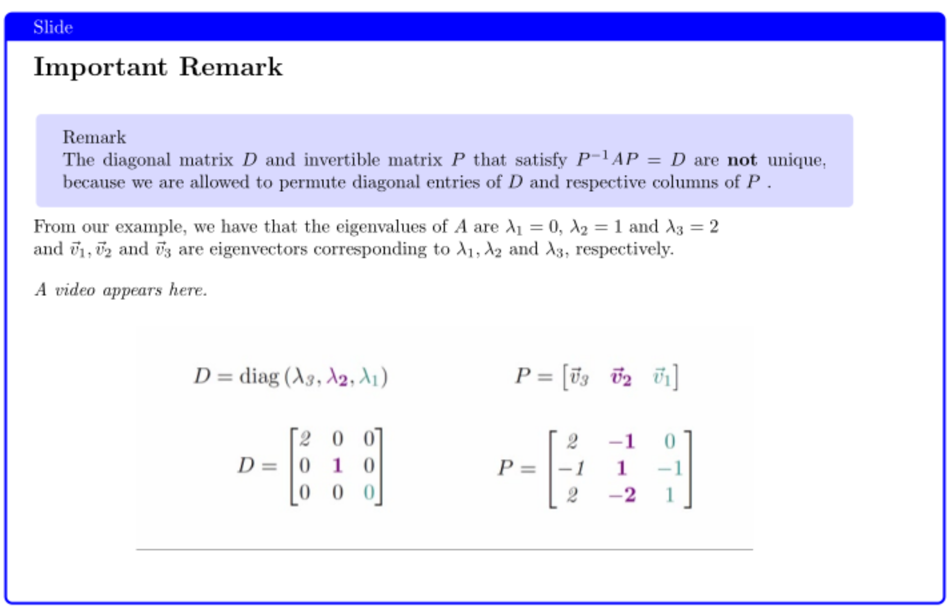
A standalone video in Mobius

In this workbook, you will see the message *"A video appears here"* followed by the **last frame** of the video.



A standalone video in the workbook

In the special cases where a video is embedded in a slide, you will also see the text *"A video appears here"* along with the **last frame** of the video in that slide

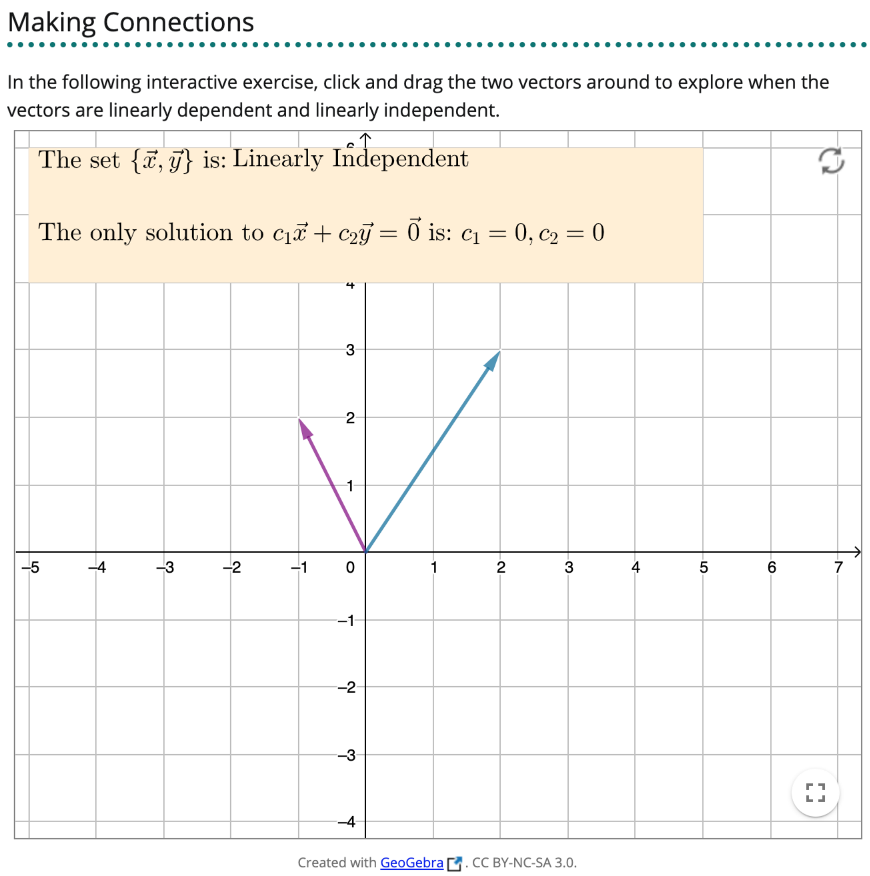


A video in a slide in the workbook

We recommend that you watch the video and write down your notes and observations in your workbook.

Making Connections

In Mobius, there are several interactive GeoGebra applets that bring course concepts to life. Each applet is presented under a heading of "Making Connections".



A GeoGebra applet in Mobius

In this workbook, the preceding text for the GeoGebra applet are printed along with a link to the GeoGebra resource.

Making Connections

In the following interactive exercise, click and drag the two vectors around to explore when the vectors are linearly dependent and linearly independent.

External resource: https://www.geogebra.org/material/iframe/id/uebb59xn/

Created with https://www.geogebra.org. CC BY-NC-SA 3.0.

A GeoGebra applet in the workbook

From the pdf, you can click on the link or type the link into your browser to visit the resource. We recommend that you take the time to explore each GeoGebra applet and take notes in your workbook.

Unit 1

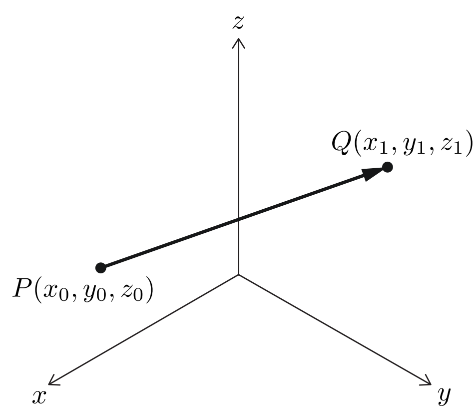
Vectors in $\mathbb{R}^n$

1.1 - Geometry, Lines, and Planes in $\mathbb{R}^n$

Geometry of $\mathbb{R}^n$

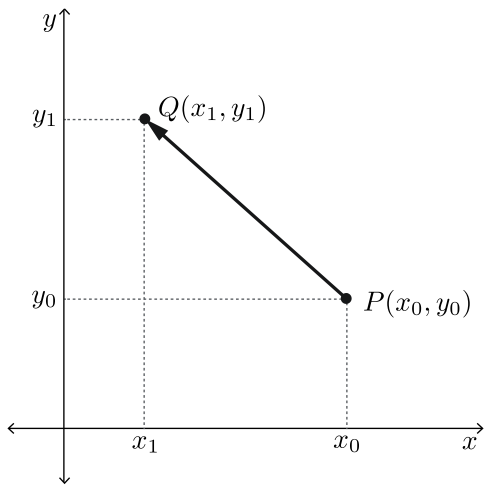
The Vector Space $\mathbb{R}^n$

You may have previously encountered the concept of **vectors** as directed line segments connecting two points on a plane or two points in 3D space.

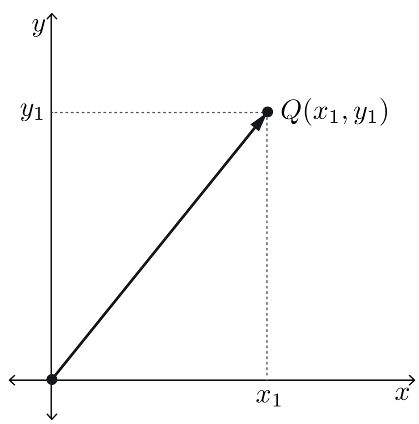


Both a plane and a 3D space are examples of **vector spaces**. In this course, we study properties of vector spaces and vectors, which are the elements of vector spaces.

Let us recall what vectors on a plane look like. Given two points, $P(x_0, y_0)$ and $Q(x_1, y_1)$, we can define a vector connecting P and Q, which points in the direction of Q.



In this course, however, we will always assume that $P(x_0, y_0)$ is the point at the origin; that is, $P(x_0, y_0) = (0, 0)$. As a consequence, our vectors will be defined only through the coordinates of the point $Q(x_1, y_1)$, as shown below.



We will denote such a vector by

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

We call the collection of all vectors of this form $\mathbb{R}^2$:

$$\mathbb{R}^2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x, y \in \mathbb{R} \right\}$$

Analogously, we can define $\mathbb{R}^3$ as the collection of all vectors connecting the point at the origin $(0, 0, 0)$ to some point $Q(x, y, z)$ in a 3D space:

$$\mathbb{R}^3 = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x, y, z \in \mathbb{R} \right\}$$

More generally, for a positive integer n, we define $\mathbb{R}^n$ as the collection of all vectors that connect the origin to a point $Q(x_1, x_2, \ldots, x_n)$.

Definition
Let n be a positive integer. The vector space $\mathbb{R}^n$ is defined by

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mid x_1, x_2, \ldots, x_n \in \mathbb{R} \right\}$$

The values $x_1, x_2, \ldots, x_n$ are called the **coordinates** of a vector $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

To distinguish the vector from its coordinates, we use the arrow symbol above a letter of the Latin alphabet.

In this case, for example, we could write $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

Exercise 1

A question appears in Mobius

Remark
In this course, we will often use geometric interpretation of $\mathbb{R}^2$, $\mathbb{R}^3$ or $\mathbb{R}^n$. However, this is not the only possible interpretation of these vector spaces. We can use vectors as tools to record data. For example, imagine a bookstore selling 1) books; 2) magazines; 3) newspapers; and 4) notebooks. They can keep track of their inventory by recording quantitative data into a vector in $\mathbb{R}^4$.

The **zero vector** in $\mathbb{R}^n$ is denoted by

$$\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

For example, in $\mathbb{R}^2$, $\vec{0}_{\mathbb{R}^2} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and in $\mathbb{R}^4$, $\vec{0}_{\mathbb{R}^4} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$. It is typical to omit the subscript as the zero vector being used is usually clear from the context.

Finally, we say that two vectors $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ are equal if and only if

$$v_1 = w_1, v_2 = w_2, \ldots, v_n = w_n$$

We will see in the next sections how to combine vectors in $\mathbb{R}^n$ to create new ones.

Addition of Vectors

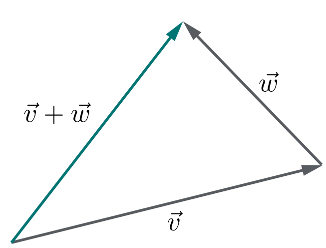
Now that we have defined $\mathbb{R}^n$ and vectors in $\mathbb{R}^n$, we will discuss what operations we can perform on them. We will start with the **addition** operation.

Definition

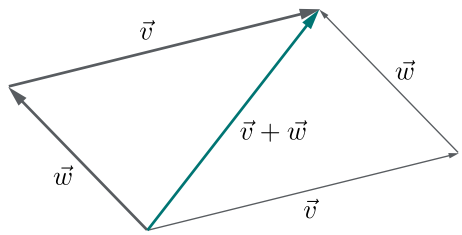
Given two vectors $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ in $\mathbb{R}^n$, we define their **sum** as

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

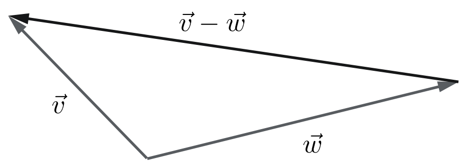
Geometrically, vector addition can be seen as adding the vectors "tip-to-tail". That is, the vector $\vec{v} + \vec{w}$ can be found by "gluing" the tail of $\vec{w}$ to the tip of $\vec{v}$



or vice-versa

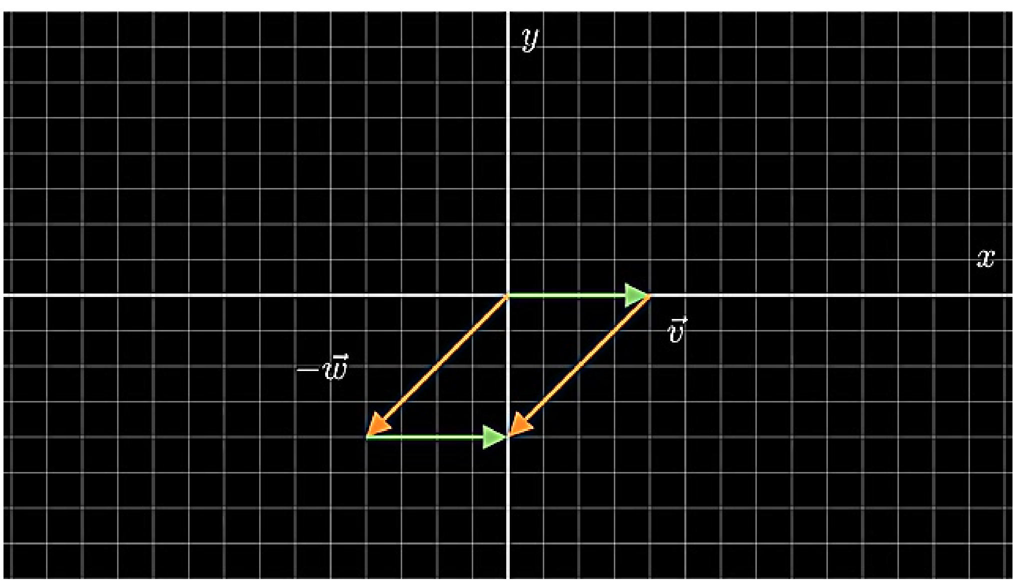


The vector from the tip of $\vec{w}$ to the tip of $\vec{v}$ is given by $\vec{v}+(-\vec{w})$, or simply $\vec{v}-\vec{w}$:



Watch the following video, which demonstrates addition and subtraction of vectors in $\mathbb{R}^2$. Note that this video has no sound.

A video appears here.



Example 1

The sum of the vectors $\begin{bmatrix} \pi \\ -1 \\ 0 \\ -1/2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ \sqrt{2} \\ -1 \\ 3/2 \end{bmatrix}$ in $\mathbb{R}^4$ can be computed as follows:

$$\begin{bmatrix} \pi \\ -1 \\ 0 \\ -1/2 \end{bmatrix} + \begin{bmatrix} 2 \\ \sqrt{2} \\ -1 \\ 3/2 \end{bmatrix} = \begin{bmatrix} \pi + 2 \\ -1 + \sqrt{2} \\ -1 \\ 1 \end{bmatrix}$$

Example 2

The vector from $A(1, 1, 1)$ to $B(2, 3, 4)$ is

$$\vec{AB} = \vec{OB} - \vec{OA} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Remark
It's important to note that in order for the sum of vectors to be defined, the vectors must have the same number of components. For example, the sum $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is not defined. While it is tempting to write $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$, these are not the same object since $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathbb{R}^2$ but $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \in \mathbb{R}^3$.

Exercise 2

A question appears in Mobius

Scalar Multiplication of Vectors

We've previously seen that we can perform addition on vectors. We will now discuss scalar multiplication of vectors.

Definition

Given a real number $c \in \mathbb{R}$ and a vector $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, we define their **scalar multiplication** as

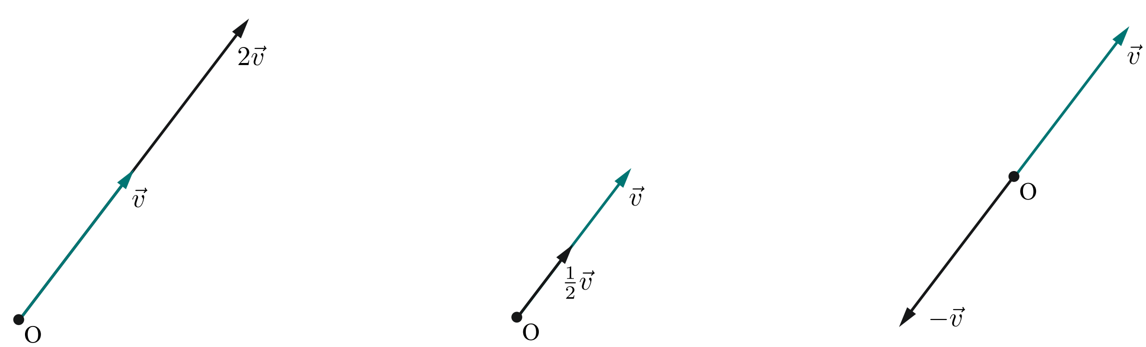
$$c \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$$

We refer to number c as a **scalar**.

Geometrically, scalar multiplication can be thought of as stretching the original vector $\vec{v}$ when $|c| \geq 1$, compressing $\vec{v}$ when $|c| < 1$ and changing the direction of $\vec{v}$ when $c < 0$. These cases are shown, respectively, in the diagrams below.

Stretching **Compressing** **Changing direction**

when $|c| \geq 1$ when $|c| < 1$ when $c < 0$



Remark
We refer to $\mathbb{R}^n$ as a **real** vector space because we choose our scalars to be **real numbers** (that is, elements of $\mathbb{R}$). In this course, we will always choose real scalars, but other choices are also possible.

Example 3

The scalar product of 3 and the vector $\begin{bmatrix} \pi \\ -1 \\ 0 \\ -1/2 \end{bmatrix}$ in $\mathbb{R}^4$ can be computed as follows:

$$3 \begin{bmatrix} \pi \\ -1 \\ 0 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 3\pi \\ -3 \\ 0 \\ -3/2 \end{bmatrix}$$

Exercise 3

A question appears in Mobius

Remark
Two non-zero vectors $\vec{v}$ and $\vec{w}$ are **parallel** if and only if they are scalar multiples of each other.

Example 4

The vectors $\vec{v} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} -4 \\ 10 \end{bmatrix}$ are parallel since $\vec{w} = -2 \cdot \vec{v}$ (or $\vec{v} = -1/2 \cdot \vec{w}$).

The vectors $\vec{v} = \begin{bmatrix} -2 \\ -3 \\ -4 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} -2 \\ -1 \\ -13 \end{bmatrix}$ are not parallel, since $\vec{v} \neq k \cdot \vec{w}$ for $k \in \mathbb{R}$. We can see this by equating components:

$$\begin{aligned} -2 &= k(-2) \Rightarrow k = 1 \\ -3 &= k(-1) \Rightarrow k = 3 \\ -4 &= k(-13) \Rightarrow k = 4/13 \end{aligned}$$

Any pair of the above equations are in contradiction; therefore, such a $k \in \mathbb{R}$ does not exist, which means that $\vec{v}$ and $\vec{w}$ are not parallel.

Exercise 4

A question appears in Mobius

Linear Combinations of Vectors

We can combine the operations of addition and scalar multiplication in the following way:

Definition
Given k scalars $c_1, c_2, \ldots, c_k \in \mathbb{R}$ and k vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in \mathbb{R}^n$, we refer to

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$$

as a **linear combination** of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$.

Example 5

The vector $\begin{bmatrix} 6 \\ -2 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in $\mathbb{R}^2$, because

$$\begin{bmatrix} 6 \\ -2 \end{bmatrix} = 3\begin{bmatrix} 2 \\ 1 \end{bmatrix} - 5\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

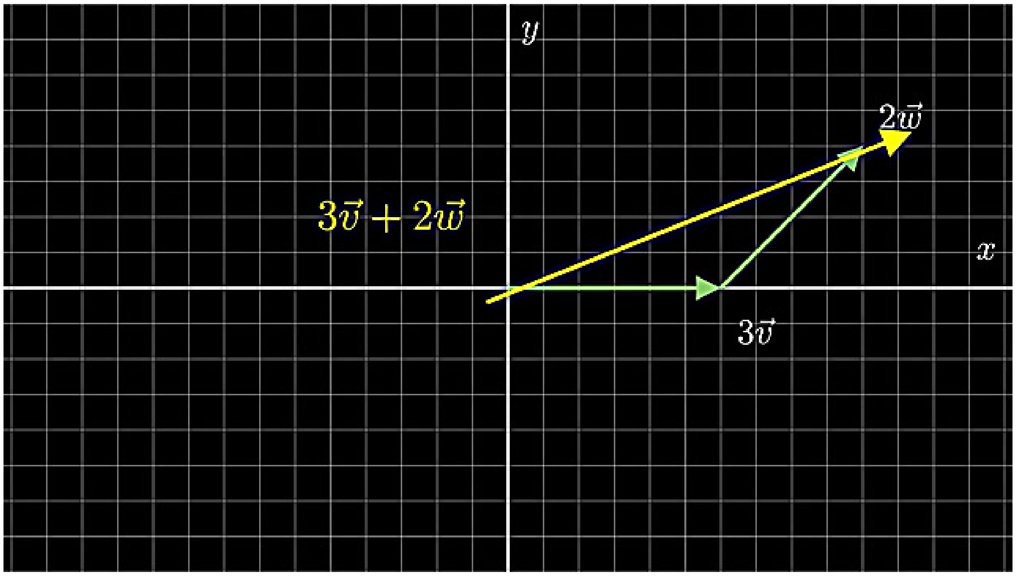
Example 6

The vector $\begin{bmatrix} -13 \\ -4 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ in $\mathbb{R}^2$, because

$$\begin{bmatrix} -13 \\ -4 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2 \\ 6 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

Watch the following video, which demonstrates linear combination of two vectors. Note that this video has no sound.

A video appears here.



Making Connections

This applet illustrates the linear combination $a\vec{u} + b\vec{v}$ of the vectors $\vec{u}$ and $\vec{v}$ for real numbers a and b.

Instructions: Change $\vec{u}$ and $\vec{v}$ by dragging their tips to a new location. Change the values of a and b by using the sliders.

External resource: https://www.geogebra.org/material/iframe/id/uwpyekrr/

Exercise 5

A question appears in Mobius

Remark
Notice that adding (or subtracting) vectors, performing scalar multiplication on a vector, and taking linear combinations of vectors in $\mathbb{R}^n$ all produce vectors in $\mathbb{R}^n$. Because of this, we say that $\mathbb{R}^n$ is **closed** under addition, scalar multiplication, and linear combinations. We will come back to these ideas later in the course.

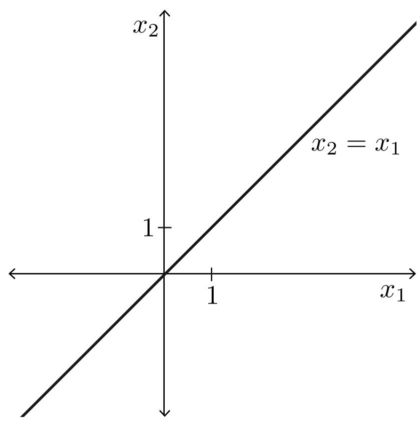
Lines in $\mathbb{R}^n$

Lines in Space

In $\mathbb{R}^2$, we define lines by equations such as

$$x_2 = mx_1 + b \quad \text{or} \quad ax_1 + bx_2 = c$$

Consider the graph of the line $x_2 = x_1$ in $\mathbb{R}^2$. The graph of the line consists of all the points (x_1, x_2) such that $x_2 = x_1$.

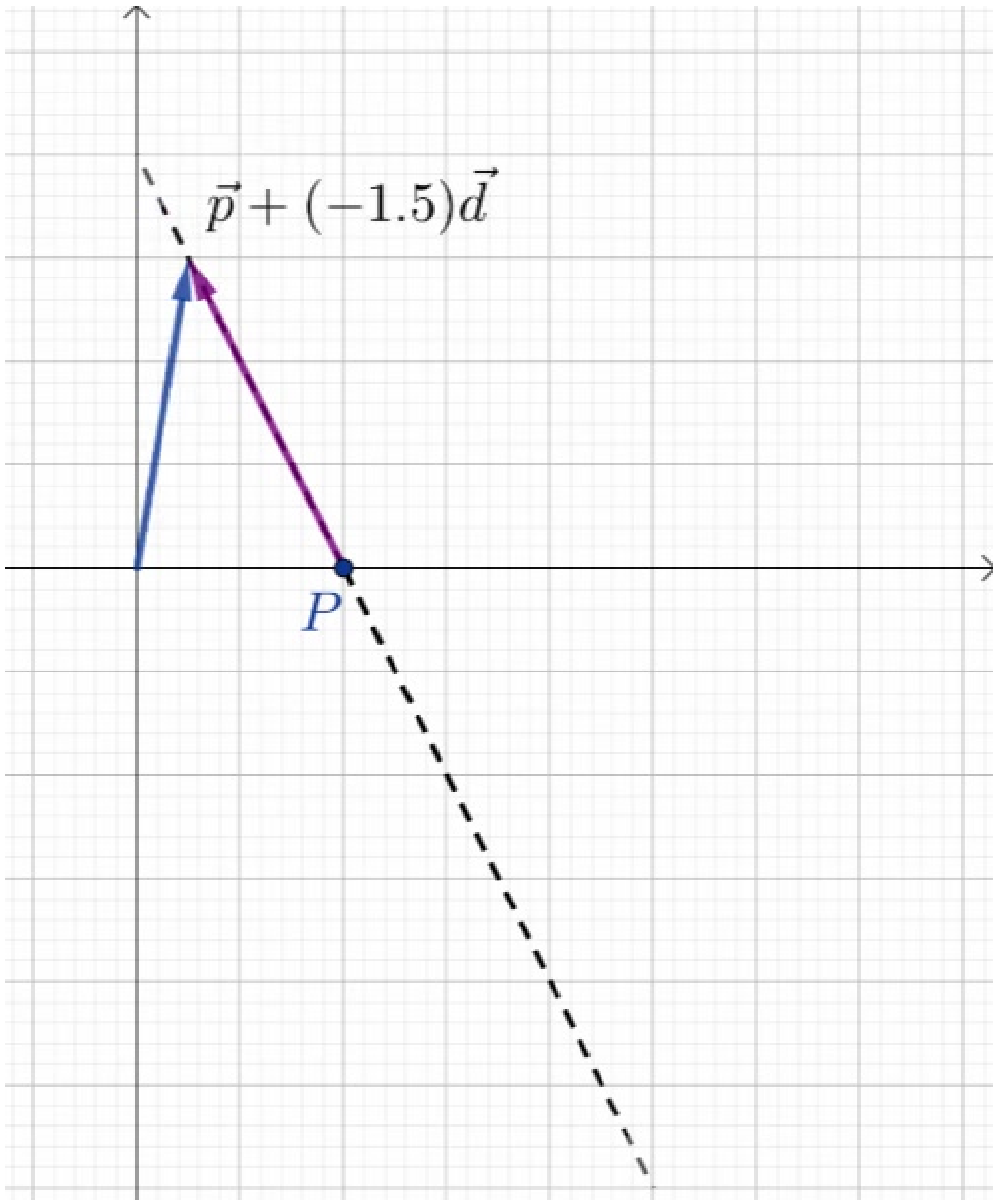


Notice that we need two things to describe a line:

1. A point P on the line
2. A vector $\vec{d}$ in the direction of the line

Let $\vec{p}$ be the vector connecting the origin O to the point P, that is, $\vec{p} = \vec{OP}$. The following animation illustrates how the vector $\vec{p} + t\vec{d}$ "draws out" the line as t changes:

A video appears here.



In $\mathbb{R}^3$, a line through a point P with direction $\vec{d}$ is given by

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \vec{p} + t \cdot \vec{d}$$

where $\vec{p}$ is the vector which connects the origin and point P, $t \in \mathbb{R}$ is a scalar, and $\vec{d} \neq \vec{0}$.

Making Connections

Instructions: Click and drag the direction vector $\vec{d}$ to find the line that passes through the point $P = (2, 2)$ and the point A. Once youve found the line, and hence the direction vector, verify your work algebraically. That is, check that the vector $OA - OP$ is a multiple of the $\vec{d}$.

External resource: https://www.geogebra.org/material/iframe/id/wn5dxser/

Example

To find the line through the points $A(1, 1, -1)$ and $B(4, 0, -3)$, we need a vector in the direction of the line and a point on the line.

The direction vector $\vec{d}$ is given by

$$\vec{d} = \vec{AB} = \vec{OB} - \vec{OA} = \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}$$

Since both A and B are points on the line, we can choose either. Choosing A, the line through the points A and B is

$$\vec{x} = \vec{OA} + t\vec{d} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + t \cdot \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}, \quad t \in \mathbb{R}$$

Or, equivalently, choosing B, the line is

$$\vec{x} = \vec{OB} + t\vec{d} = \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix} + t \cdot \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}, \quad t \in \mathbb{R}$$

In the above example, we could have used **any** known point on the line and **any** non-zero scalar multiple of the direction vector. This means that there are infinitely many ways to represent a line.

Vector and Parametric Equations of a Line

In this section, we will see two equivalent ways of describing a line in $\mathbb{R}^n$.

Definition
Let $\vec{p}, \vec{d} \in \mathbb{R}^n$ and $t \in \mathbb{R}$. The **vector equation** of the line through $\vec{p}$ with direction vector $\vec{d}$ is

$$\vec{x} = \vec{p} + t \cdot \vec{d}$$

Thus, for each $t \in \mathbb{R}$, the vector equation returns the position vector of a point on the line.

We sometimes like to expand the vector equation:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} + t \cdot \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = \begin{bmatrix} p_1 + t \cdot d_1 \\ p_2 + t \cdot d_2 \\ \vdots \\ p_n + t \cdot d_n \end{bmatrix}$$

Equating entries gives us the **parametric equations** of the line:

Definition

Let $\vec{p}, \vec{d} \in \mathbb{R}^n$ and $t \in \mathbb{R}$. The **parametric equations** of the line through $\vec{p}$ with direction vector $\vec{d}$ are

$$\begin{aligned} x_1 &= p_1 + td_1 \\ x_2 &= p_2 + td_2 \\ \vdots &= \vdots \\ x_n &= p_n + td_n \end{aligned}$$

Thus, for each $t \in \mathbb{R}$, the parametric equations give the x_k-coordinates of a point on the line for $k = 1, 2, \ldots, n$.

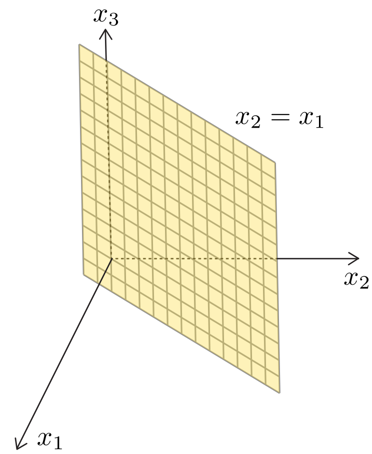
Exercise

A question appears in Mobius

Planes in $\mathbb{R}^n$

Planes in Space

Recall the equation $x_2 = x_1$ which we saw earlier in $\mathbb{R}^2$. Let's now consider this equation in $\mathbb{R}^3$. We are considering all the points (x_1, x_2, x_3) such that $x_2 = x_1$. Since there is no restriction on x_3, this means that x_3 can be any real number. Hence the equation $x_2 = x_1$ represents a plane in $\mathbb{R}^3$.



In order to describe a plane, we need three things:

1. A point P on the plane
2. A non-zero vector $\vec{u}$ lying in the plane
3. A second non-zero $\vec{v}$ vector lying in the plane which is not parallel to $\vec{u}$

In $\mathbb{R}^3$, a plane through a point P containing vectors $\vec{u}$ and $\vec{v}$ which are non-zero and not parallel is given by

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \vec{p} + s\vec{u} + t\vec{v} \quad s, t, \in \mathbb{R}$$

where $\vec{p} = \vec{OP}$ is the vector connecting the origin to the point p and $\vec{u}, \vec{v}$ lie on the plane.

Making Connections

Instructions: Click and drag the tips of vectors $\vec{u}$ and $\vec{v}$ to find the plane that passes through the point $P = (0, 1, 1)$ and the point $A = (-1, 2, 2)$. Once youve found a set of vectors $\vec{u}$ and $\vec{v}$, verify your work algebraically. Can you find more than one way of doing this? Why or why not?

External resource: https://www.geogebra.org/material/iframe/id/vk9t2mvc/

Example 1

To find the equation of the plane containing the points $A(1, 1, 1)$, $B(1, 2, 3)$, and $C(-1, 1, 2)$ we need two non-zero and non-parallel vectors that lie on the plane and a point on the plane.

We have

$$\vec{AB} = \vec{OB} - \vec{OA} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

and

$$\vec{AC} = \vec{OC} - \vec{OA} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

Note that $\vec{AB}$ and $\vec{AC}$ are non-zero and not parallel, so we have found the two vectors we need. As for the point on the plane, we choose the point $\vec{A} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, which is given to us. Thus, the equation

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + s \cdot \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \quad s, t, \in \mathbb{R}$$

is an equation which describes our plane.

Vector and Parametric Equations of a Plane

Just as we did with the vector and parametric equations of lines, we can define vector and parametric equations of a plane.

Definition
Given a point P on a plane in $\mathbb{R}^n$ and a pair of non-zero and not parallel vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ lying in the plane, the **vector equation** of the plane is given by

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \vec{p} + s\vec{u} + t\vec{v}, \quad s, t, \in \mathbb{R}$$

where $\vec{p} = \vec{OP}$.

As with lines, this representation is not unique: it depends on our choice of $\vec{p}$, $\vec{u}$, and $\vec{v}$.

We can also define parametric equations for planes:

Definition
Given a vector equation of a plane $\vec{x} = \vec{p} + s\vec{u} + t\vec{v}$, the corresponding **parametric equations** are

$$\begin{cases} x_1 = p_1 + s \cdot u_1 + t \cdot v_1 \\ x_2 = p_2 + s \cdot u_2 + t \cdot v_2 \\ \vdots \\ x_n = p_n + s \cdot u_n + t \cdot v_n \end{cases}$$

Remark
If the two non-zero vectors $\vec{u}, \vec{v}$ lying in the plane are parallel, i.e., $\vec{u} = k\vec{v}$, then the vector equation

$$\vec{x} = \vec{p} + s\vec{u} + t\vec{v} = \vec{p} + s(k\vec{v}) + t\vec{v} = \vec{p} + (sk + t)\vec{v}$$

describes a line, not a plane.

Example 2

To find a vector equation of the plane containing the point $P(1,-1,-2)$ and the line

$$\vec{x} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} + r \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}, \quad r \in \mathbb{R}$$

we notice that we have been given a point on the plane as well as one of the vectors on the plane.

The line equation $\vec{x} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} + r \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ tells us that $\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$ is another point on the plane and that $\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ is a direction vector. We need to use the points $\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$ and $\vec{p} = \vec{OP}$ to find a second non-zero vector in the plane which is not parallel to $\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$. We have

$$\vec{v} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -4 \\ -1 \end{bmatrix}$$

This vector is non-zero and not parallel to $\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$, so a vector equation for the plane is

$$\vec{x} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} + s \cdot \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} + t \cdot \begin{bmatrix} 0 \\ -4 \\ -1 \end{bmatrix} \quad s, t, \in \mathbb{R}$$

Exercise

A question appears in Mobius

1.2 - Norm, Dot Product, and Cross Product

Norm of a Vector

Definition and Properties of the Norm

Length of a Vector

One thing we often want to know about a vector is its length or norm. We define it as follows:

Definition

The **norm** (or magnitude or length) of a vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$ is

$$\|\vec{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}$$

Example 1

Let $\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathbb{R}^2$. Find the norm of $\vec{v}$.

Solution

The norm of $\vec{v}$ is

$$\begin{aligned} \|\vec{v}\| &= \sqrt{1^2 + 2^2} \\ &= \sqrt{1+4} \\ &= \sqrt{5} \end{aligned}$$

Example 2

Let $\vec{w} = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 3 \end{bmatrix} \in \mathbb{R}^4$. Find the norm of $\vec{w}$.

Solution

The norm of $\vec{w}$ is

$$\begin{aligned}\|\vec{w}\| &= \sqrt{1^2+2^2+(-2)^2+3^2} \\ &= \sqrt{1+4+4+9} \\ &= \sqrt{18}\end{aligned}$$

Exercise 1

A question appears in Mobius

Distance Between Two Vectors

In addition to determining the length of a vector, we also use the norm to find the distance between two vectors.

Definition

Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$. The **distance** between $\vec{v}$ and $\vec{w}$ is given by

$$\|\vec{v}-\vec{w}\| = \sqrt{(v_1-w_1)^2+\cdots+(v_n-w_n)^2}$$

Example 3

Let $\vec{v} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$. The distance from $\vec{v}$ to $\vec{w}$ is

$$\begin{aligned}\|\vec{v}-\vec{w}\| &= \sqrt{(3-1)^2+(2-(-1))^2+(1-2)^2} \\ &= \sqrt{2^2+3^2+(-1)^2} \\ &= \sqrt{4+9+1} \\ &= \sqrt{14}\end{aligned}$$

Exercise 2

A question appears in Mobius

In the following proposition, we list some useful properties satisfied by the norm.

Proposition 1: Norm of a Vector
Let $\vec{v}, \vec{w} \in \mathbb{R}^n$ and $k \in \mathbb{R}$

1. $\|\vec{v}\| \geq 0$ with equality if and only if $\vec{v} = \vec{0}$
2. $\|k\vec{v}\| = |k|\|\vec{v}\|$
3. $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$; this is called the **triangle inequality**

- Property 1 means that the norm of a vector is always non-negative; this agrees with our intuition that lengths should not be negative.
- Property 2 tells us that taking the norm of a vector which has been scaled is the same as taking the norm of the vector, then multiplying it by the absolute value of the scaling factor.
- Property 3 tells us that the norm of the sum of two vectors is less than or equal to the sum of the norms of these vectors. The triangle inequality gets its name from the fact that for any triangle, the sum of the length of any two sides must be less than or equal to the length of the third side.

Unit Vector

There is a special type of vector called a unit vector.

Definition
A vector $\vec{v} \in \mathbb{R}^n$ is a **unit vector** if $\|\vec{v}\| = 1$.

Example 4

- $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is a unit vector since $\|\vec{v}\| = \sqrt{1^2 + 0^2} = 1$.

- $\vec{w} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$ is a unit vector since $\|\vec{w}\| = \sqrt{(1/\sqrt{3})^2 + (1/\sqrt{3})^2 + (1/\sqrt{3})^2} = \sqrt{1/3 + 1/3 + 1/3} = 1.$

We can make an arbitrary non-zero vector $\vec{v} \in \mathbb{R}^n$ into a unit vector by dividing $\vec{v}$ by its norm, $\|\vec{v}\|$. This is called **normalizing** $\vec{v}$. The normalized vector is a unit vector in the direction of $\vec{v}$ since

$$\left\| \frac{1}{\|\vec{v}\|}\vec{v} \right\| = \left| \frac{1}{\|\vec{v}\|} \right| \|\vec{v}\| = \frac{1}{\|\vec{v}\|} \|\vec{v}\| = 1$$

and the scalar $\dfrac{1}{\|\vec{v}\|}$ is positive.

Example 5

Let $\vec{v} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$. Find a unit vector in the direction of $\vec{v}$.

Solution

A unit vector in the direction of $\vec{v}$ is given by $\dfrac{1}{\|\vec{v}\|}\vec{v}$ where

$$\begin{aligned} \|\vec{v}\| &= \sqrt{4^2 + 5^2 + 6^2} \\ &= \sqrt{16 + 25 + 36} \\ &= \sqrt{77} \end{aligned}$$

Therefore $\dfrac{1}{\sqrt{77}} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 4/\sqrt{77} \\ 5/\sqrt{77} \\ 6/\sqrt{77} \end{bmatrix}$ is a unit vector in the direction of $\vec{v}$.

Exercise 3

A question appears in Mobius

Dot Product

Definition and Properties of Dot Product

Definition

Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$. The **dot product** of $\vec{v}$ and $\vec{w}$ is

$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = \sum_{i=1}^{n} v_i w_i$$

Example 1

Let $\vec{v} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} -3 \\ -4 \\ 5 \end{bmatrix}$.

Then,

$$\begin{aligned} \vec{v} \cdot \vec{w} &= \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ -4 \\ 5 \end{bmatrix} \\ &= 1(-3) + 1(-4) + 2(5) \\ &= -3 - 4 + 10 = 3 \end{aligned}$$

Exercise 1

A question appears in Mobius

The dot product operation has several interesting and useful properties, which we summarize in the following proposition.

Proposition 2: Dot Product
Let $\vec{v}$, $\vec{u}$, $\vec{w} \in \mathbb{R}^n$, and $k \in \mathbb{R}$.

1. $\vec{v} \cdot \vec{w} \in \mathbb{R}$; the dot product always outputs a scalar
2. $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$
3. $\vec{v} \cdot \vec{0} = 0$
4. $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$
5. $(k\vec{v}) \cdot \vec{w} = k(\vec{v} \cdot \vec{w}) = \vec{v} \cdot (k\vec{w})$
6. $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$ and $\vec{u} \cdot (\vec{v} - \vec{w}) = \vec{u} \cdot \vec{v} - \vec{u} \cdot \vec{w}$
7. $|\vec{v} \cdot \vec{u}| \leq \|\vec{v}\|\|\vec{u}\|$; this is called the **Cauchy-Schwarz Inequality** and states that the size of the dot product of $\vec{v}$ and $\vec{u}$ cannot exceed the product of their norms.

Proof

Let us now prove property 2 of the dot product, $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$.

Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$.

Then,

$$\begin{aligned} \vec{v} \cdot \vec{w} &= v_1 w_1 + \cdots + v_n w_n && \text{by definition of dot product} \\ &= w_1 v_1 + \cdots + w_n v_n && \text{by properties of real numbers} \\ &= \vec{w} \cdot \vec{v} && \text{by definition of dot product} \end{aligned}$$

□

Exercise 2

Prove property 4 of the dot product, $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$.

A question appears in Mobius

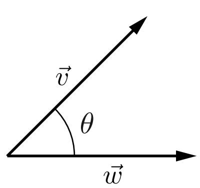
Exercise 3

Prove property 5 of the dot product, $(k\vec{v}) \cdot \vec{w} = k(\vec{v} \cdot \vec{w}) = \vec{v} \cdot (k\vec{w})$.

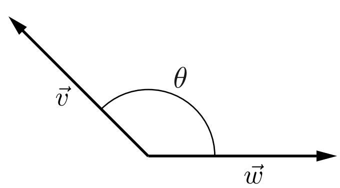
A question appears in Mobius

Orthogonality

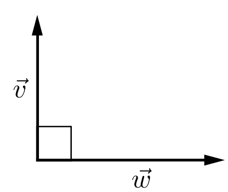
Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ be non-zero vectors in $\mathbb{R}^n$ having a common tail. The vectors determine a unique angle θ with $0 \leq \theta \leq \pi$. This angle can be **acute** $(0 \leq \theta < \pi/2)$, **obtuse** $(\pi/2 < \theta \leq \pi)$, or **right** $(\theta = \pi/2)$:



Acute: $0 \leq \theta < \pi/2$



Obtuse: $\pi/2 < \theta \leq \pi$



Right: $\theta = \pi/2$

The angle θ satisfies $\vec{v} \cdot \vec{w} = \|\vec{v}\|\|\vec{w}\| \cos(\theta)$. Furthermore, since $\vec{v}, \vec{w} \neq \vec{0}$, $\|\vec{v}\|\|\vec{w}\| > 0$, we have $\cos(\theta) = \dfrac{\vec{v} \cdot \vec{w}}{\|\vec{v}\|\|\vec{w}\|}$. We can use this information to compute angles between vectors.

Example 2

The angle between $\vec{v} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$ is

$$\begin{aligned} \theta &= \cos^{-1}\left(\frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\|\vec{w}\|}\right) \\ &= \cos^{-1}\left(\frac{2(1)+1(-1)-1(-2)}{\sqrt{2^2+1^2+(-1)^2}\sqrt{1^2+(-1)^2+(-2)^2}}\right) \\ &= \cos^{-1}\left(\frac{3}{\sqrt{6}\sqrt{6}}\right) \\ &= \cos^{-1}\left(\frac{1}{2}\right) \\ &= \pi/3 \end{aligned}$$

Note that the sign of $\cos(\theta)$ is determined by the sign of $\vec{v}\cdot\vec{w}$:

- $\vec{v}\cdot\vec{w} > 0 \Leftrightarrow 0 \le \theta < \pi/2$; the angle is acute
- $\vec{v}\cdot\vec{w} = 0 \Leftrightarrow \theta = \pi/2$; the angle is right
- $\vec{v}\cdot\vec{w} < 0 \Leftrightarrow \pi/2 < \theta \le \pi$; the angle is obtuse

Example 3

For $\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$,

$$\vec{v}\cdot\vec{w} = 1(6) + 2(-2) = 2 > 0$$

Therefore $\vec{v}$ and $\vec{w}$ determine an acute angle.

Exercise 4

A question appears in Mobius

You may have noticed that we started this section by considering only non-zero vectors. Let's see what happens for $\vec{0} \in \mathbb{R}^n$.

Since $\vec{v} \cdot \vec{0} = 0$ for every $\vec{v} \in \mathbb{R}^n$, we define $\vec{0} \in \mathbb{R}^n$ to be orthogonal to every $\vec{v} \in \mathbb{R}^n$; however, the angle determined by $\vec{0}$ and $\vec{v}$, $\cos(\theta) = \dfrac{\vec{v} \cdot \vec{0}}{\|\vec{v}\|\|\vec{0}\|}$, is not defined. Our angle calculation therefore requires that we use non-zero vectors.

The concept of vectors being orthogonal is an important one that we will see throughout the course.

Definition
Two vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$ are **orthogonal** if and only if $\vec{v} \cdot \vec{w} = 0$.

Example 4

The vectors $\vec{v} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 3 \\ 2 \\ -1/2 \end{bmatrix}$ are orthogonal since

$$\begin{aligned} \vec{v} \cdot \vec{w} &= (1)(3) + (-1)(2) + (2)(-1/2) \\ &= 3 - 2 - 1 \\ &= 0 \end{aligned}$$

Making Connections

Instructions: Click and drag the vector $\vec{v}$ so that it becomes orthogonal to the vector $\vec{u}$. Once youve found an appropriate vector $\vec{v}$, verify your work algebraically.

External resource: https://www.geogebra.org/material/iframe/id/mbqwhtda/

Exercise 5

A question appears in Mobius

Cross Product

Definition and Properties of the Cross Product

Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ be vectors in $\mathbb{R}^3$. The **cross product** of $\vec{v}$ and $\vec{w}$ is

$$\vec{v} \times \vec{w} = \begin{bmatrix} v_2 w_3 - w_2 v_3 \\ -(v_1 w_3 - w_1 v_3) \\ v_1 w_2 - w_1 v_2 \end{bmatrix}$$

It's very important to note that the cross product is an operation that only works for vectors in $\mathbb{R}^3$.

Remark
The cross product of $\vec{v}$ and $\vec{w}$ always gives a vector in $\mathbb{R}^3$. This is in contrast with the dot product, which always outputs a scalar.

Example 1

Let $\vec{v} = \begin{bmatrix} 1 \\ 6 \\ 3 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}$.

Then, the cross product is:

$$\begin{aligned} \vec{v} \times \vec{w} &= \begin{bmatrix} 6(2) - 3(3) \\ -(1(2) - (-1)(3)) \\ 1(3) - (-1)(6) \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ -5 \\ 9 \end{bmatrix} \end{aligned}$$

Exercise 1

A question appears in Mobius

The cross product operation has several interesting and useful properties, which we summarize in the following proposition:

Proposition 3: Cross Product
Let $\vec{v}$, $\vec{u}$, $\vec{w} \in \mathbb{R}^3$, and $k \in \mathbb{R}$.

1. $\vec{v} \times \vec{u} \in \mathbb{R}^3$
2. $\vec{v} \times \vec{u}$ is orthogonal to both $\vec{v}$ and $\vec{u}$
3. $\vec{v} \times \vec{0} = \vec{0} = \vec{0} \times \vec{v}$
4. $\vec{v} \times \vec{v} = \vec{0}$
5. $\vec{v} \times \vec{u} = -(\vec{u} \times \vec{v})$
6. $k(\vec{v} \times \vec{u}) = (k\vec{v}) \times \vec{u} = \vec{v} \times (k\vec{u})$
7. $\vec{v} \times (\vec{u} + \vec{w}) = (\vec{v} \times \vec{u}) + (\vec{v} \times \vec{w})$ and $\vec{v} \times (\vec{u} - \vec{w}) = (\vec{v} \times \vec{u}) - (\vec{v} \times \vec{w})$
8. $(\vec{v} + \vec{u}) \times \vec{w} = (\vec{v} \times \vec{w}) + (\vec{u} \times \vec{w})$ and $(\vec{v} - \vec{u}) \times \vec{w} = (\vec{v} \times \vec{w}) - (\vec{u} \times \vec{w})$

Property 2 is particularly relevant for us as we are often looking for orthogonal vectors. The property tells us that to find a vector in $\mathbb{R}^3$ which is orthogonal to both vectors $\vec{v}$ and $\vec{u}$, we only need to calculate the vector $\vec{v} \times \vec{u}$.

From some of the other properties stated in the above proposition, it follows that if non-zero vectors $\vec{v}$ and $\vec{u}$ are parallel, then $\vec{v} \times \vec{u} = \vec{0}$. Let's prove this:

Suppose that $\vec{v}$ and $\vec{u}$ are parallel, that is, $\vec{v} = k\vec{u}$.

Then

$$\begin{aligned}
\vec{v} \times \vec{u} &= (k\vec{u}) \times \vec{u} && \text{by hypothesis} \\
&= k(\vec{u} \times \vec{u}) && \text{by Property 6} \\
&= k\,\vec{0} && \text{by Property 4} \\
&= \vec{0}
\end{aligned}$$

Exercise 2

Prove property 5 for cross products.

A question appears in Mobius

Exercise 3

A question appears in Mobius

Cross Product and Orthogonality

Let's take a closer look at Property 2 of the proposition about properties of the cross product:

Cross Product: Proposition - Property 2

Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \in \mathbb{R}^3$. Then $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$.

Proof

Let us now prove this property. In order to show that $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$, we must show that $\vec{v} \cdot (\vec{v} \times \vec{w}) = 0$ and that $\vec{w} \cdot (\vec{v} \times \vec{w}) = 0$.

We have

$$\begin{aligned}
\vec{v}\cdot(\vec{v}\times\vec{w}) &= \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \cdot \begin{bmatrix} v_2w_3 - w_2v_3 \\ -(v_1w_3 - w_1v_3) \\ v_1w_2 - w_1v_2 \end{bmatrix} && \text{by definition of cross product} \\
&= v_1(v_2w_3 - w_2v_3) - v_2(v_1w_3 - w_1v_3) + v_3(v_1w_2 - w_1v_2) && \text{by definition of dot product} \\
&= v_1v_2w_3 - v_1v_3w_2 - v_1v_2w_3 + v_2v_3w_1 + v_1v_3w_2 - v_2v_3w_1 \\
&= 0.
\end{aligned}$$

Therefore $\vec{v}\times\vec{w}$ is orthogonal to $\vec{v}$. The proof of $\vec{w}\cdot(\vec{v}\times\vec{w}) = 0$ proceeds analogously.

□

Making Connections

This applet illustrates the vectors $\vec{u}$, $\vec{v}$ and their cross product, $\vec{u}\times\vec{v}$.

Instructions: Click and drag the tips of the vectors $\vec{u}$ and $\vec{v}$ and observe how their cross product changes. Notice how $\vec{u}\times\vec{v}$ is always orthogonal to both $\vec{v}$ and $\vec{w}$.

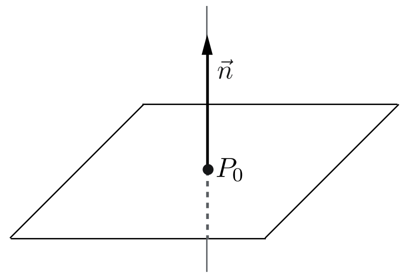
External resource: https://www.geogebra.org/material/iframe/id/fzrctcrf/

Exercise 4

A question appears in Mobius

Scalar Equation of a Plane in $\mathbb{R}^3$

Given any point P_0 on a plane in $\mathbb{R}^3$, there is a unique line through that point that is perpendicular (orthogonal) to the plane.



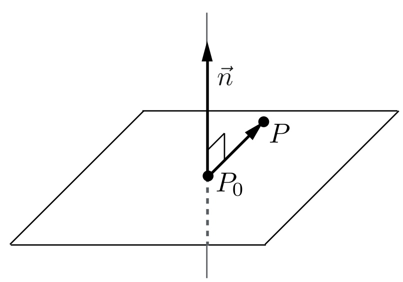
A non-zero vector $\vec{n} \in \mathbb{R}^3$ is a **normal vector** for a plane if $\vec{n}$ is orthogonal to every vector lying in the plane.

Remark
For any plane, the normal vector $\vec{n}$ through P_0 is not unique: any non-zero scalar multiple of $\vec{n}$ will also be a normal vector passing through P_0 for the plane.

We can use the normal vector to find the **scalar equation** of a plane in $\mathbb{R}^3$. Let's see how this is done.

Suppose that $P_0(a, b, c)$ is a given point on a plane with normal vector $\vec{n} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$. Let $P(x_1, x_2, x_3)$ be any point. Then, P lies on the plane if and only if $\vec{P_0P}$, the vector from P_0 to P, also lies on the plane. Equivalently, as shown in the image below, $\vec{n}$ is orthogonal to $\vec{P_0P}$; that is, $\vec{n} \cdot \vec{P_0P} = 0$. Rearranging this equality, we find that

$$\vec{n} \cdot \vec{P_0P} = 0 \iff \vec{n} \cdot (\vec{OP} - \vec{OP_0}) = 0 \iff \vec{n} \cdot \vec{OP} = \vec{n} \cdot \vec{OP_0}.$$



Expanding the last equality, we have

$$\begin{gathered} \vec{n} \cdot \vec{OP} = \vec{n} \cdot \vec{OP_0} \\ \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ n_1x_1 + n_2x_2 + n_3x_3 = n_1a + n_2b + n_3c \end{gathered}$$

Definition
Since n_1, n_2, n_3, a, b, c are given, we let $d = n_1a + n_2b + n_3c$ and say that the **scalar equation for the plane** is

$$n_1x_1 + n_2x_2 + n_3x_3 = d$$

Remark
It is also common to write the scalar equation in the form

$$n_1(x_1 - a) + n_2(x_2 - b) + n_3(x_3 - c) = 0$$

Example 2

To find a scalar equation of the plane containing the points $A(3,1,2), B(1,2,3), C(-2,1,3)$, we begin by calculating $\vec{AB}$ and $\vec{AC}$:

$$\vec{AB} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$
$$\vec{AC} = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}$$

Since $\vec{AB}$ and $\vec{AC}$ are nonzero and not parallel, we can take

$$\vec{n} = \vec{AB} \times \vec{AC} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix}$$

as a normal vector. Now, using the scalar equation of a plane

$$n_1x_1 + n_2x_2 + n_3x_3 = d$$

we substitute $n_1 = 1, n_2 = -3, n_3 = 5$to give

$$1x_1 - 3x_2 + 5x_3 = d$$

We then substitute point $A(3,1,2)$to find d :

$$\begin{aligned} 1(3) - 3(1) + 5(2) &= d \\ 10 &= d \end{aligned}$$

Thus, a scalar equation for the plane is

$$1x_1 - 3x_2 + 5x_3 = 10$$

We could also have used points B or C to get the same scalar equation.

Remark
Note that the scalar equation for a plane is not unique since any scalar multiple of a normal vector is also a normal vector. For example, $2x_1 - 6x_2 + 10x_3 = 20$ is also a scalar equation of the same plane.

Exercise 5

A question appears in Mobius

Conversion Between Vector and Scalar Equations of a Plane

Scalar Equation to Vector Equation

Lets now look at the vector equation of a plane in $\mathbb{R}^3$; specifically, how to convert from a scalar equation of a plane to a vector equation of a plane.

A slideshow appears in Mobius.

Slide

Example 3

Convert the scalar equation $x_1 + 2x_2 - 3x_3 + 6 = 0$ into a vector equation.
Solution
Recall that a vector equation of a plane in $\mathbb{R}^3$ is of the form

$$\vec{x} = \vec{p} + s\vec{u} + t\vec{v}, \quad s, t \in \mathbb{R}$$

where

- $\vec{p}$ is a vector connecting the origin to a point on the plane;
- $\vec{u}$ and $\vec{v}$ are non-parallel, non-zero vectors on the plane, that are orthogonal to the normal line of the plane.

To find the vector equation of the plane, we need to find $\vec{p}$, $\vec{u}$, and $\vec{v}$.

Slide

Example 3

Convert the scalar equation $x_1 + 2x_2 - 3x_3 + 6 = 0$ into a vector equation.
Solution
Start by finding a vector $\vec{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$ that connects the origin to a point on the plane.
We choose $\vec{p}$ arbitrarily so that the scalar equation $p_1 + 2p_2 - 3p_3 + 6 = 0$ is satisfied.
Set $p_1 = 0$ and $p_2 = 0$. Then,

$$0 + 0 - 3p_3 + 6 = 0 \quad \Rightarrow \quad p_3 = 2$$

So we let $\vec{p} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$ be the vector connecting the origin to a point on the plane.

Remark
We could have picked **any** vector $\vec{p}$ which connects the origin to an arbitrary point on the plane.

Slide

Example 3

Convert the scalar equation $x_1 + 2x_2 - 3x_3 + 6 = 0$ into a vector equation.
Solution
Next, find a normal vector to the plane, $\vec{n}$, by using the coefficients of x_1, x_2, x_3 in the given scalar equation.

$$\vec{n} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$$

We now use n to find a non-zero vector $\vec{u}$. The vector $\vec{u}$ must be orthogonal to $\vec{n}$, so $\vec{n} \cdot \vec{u} = 0$. Therefore,

$$\vec{n} \cdot \vec{u} = u_1 + 2u_2 - 3u_3 = 0$$

Setting $u_1 = 1$, $u_2 = 1$, we find that $u_3 = 1$, so $\vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Slide

Example 3

Convert the scalar equation $x_1 + 2x_2 - 3x_3 + 6 = 0$ into a vector equation.
Solution
We now find a non-zero vector $\vec{v}$, which is not parallel to $\vec{u} = \begin{bmatrix}1\\1\\1\end{bmatrix}$, and is orthogonal to our normal vector $\vec{n} = \begin{bmatrix}1\\2\\-3\end{bmatrix}$. So, $\vec{n} \cdot \vec{v} = 0$. Therefore,

$$0 = \vec{n} \cdot \vec{v} = v_1 + 2v_2 - 3v_3$$

Setting $v_1 = 1$ and $v_2 = 0$, we find $v_3 = 1/3$. Thus, $\vec{v} = \begin{bmatrix}1\\0\\1/3\end{bmatrix}$.

Slide

Example 3

Convert the scalar equation $x_1 + 2x_2 - 3x_3 + 6 = 0$ into a vector equation.
Solution
We now have everything we need:

- $\vec{p}$: a vector connecting the origin to a point on the plane. We found $\vec{p} = \begin{bmatrix}0\\0\\2\end{bmatrix}$.
- $\vec{u}, \vec{v}$: non-parallel, non-zero vectors that are orthogonal to $\vec{n} = \begin{bmatrix}1\\2\\-3\end{bmatrix}$. We found $\vec{u} = \begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\vec{v} = \begin{bmatrix}1\\0\\1/3\end{bmatrix}$.

The vector equation of a plane in $\mathbb{R}^3$ is of the form

$$\vec{x} = \vec{p} + s\vec{u} + t\vec{v}$$

Substituting the values of $\vec{p}, \vec{u}, \vec{v}$, we have

$$\vec{x} = \begin{bmatrix}0\\0\\2\end{bmatrix} + s\begin{bmatrix}1\\1\\1\end{bmatrix} + t\begin{bmatrix}1\\0\\1/3\end{bmatrix}, \quad s, t \in \mathbb{R}$$

as a vector equation of our plane.

Vector Equation to Scalar Equation

In the next example, we will look at how to convert a vector equation of a plane to a scalar equation of the same plane.

Example 4

Convert the vector equation of a plane

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} + s \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix} + t \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} \quad , \quad s, t \in \mathbb{R}$$

into a scalar equation.

Solution

Recall that a scalar equation of a plane in $\mathbb{R}^3$ is given by

$$\vec{x} \cdot \vec{n} = \vec{p} \cdot \vec{n}$$

where

- $\vec{n}$ is a normal vector to the plane; and
- $\vec{p}$ is a vector connecting the origin to a point on the plane.

Since we are given a vector equation in the form of $\vec{x} = \vec{p} + t\,\vec{u} + s\,\vec{v}$, we can find a normal vector $\vec{n}$ to the plane by taking the cross product of $\vec{u}$ and $\vec{v}$:

$$\vec{n} = \vec{u} \times \vec{v} = \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix} \times \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 17 \\ -8 \\ 11 \end{bmatrix}$$

Then, the scalar equation for the plane is $\vec{x} \cdot \vec{n} = \vec{p} \cdot \vec{n}$, where $\vec{p} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$:

$$17\,x_1 - 8\,x_2 + 11\,x_3 = 17(1) - 8(2) + 11(2)$$

Simplifying, we obtain the scalar equation

$$17\,x_1 - 8\,x_2 + 11\,x_3 = 23$$

1.3 - Vector Projections

Projection onto a Vector

Intuition for Projection onto a Vector

In the previous lessons, we have learned how to find a linear combination of vectors. Now, we will reverse this process. Given any vector, we write it as a linear combination of two other vectors.

A slideshow appears in Mobius.

Slide

Intuition for Vector Projections

When working in $\mathbb{R}^2$, we can express a vector as a linear combination of two other vectors. For example,

$$\begin{aligned}\vec{u} &= \begin{bmatrix}x_1\\x_2\end{bmatrix}\\ &= \begin{bmatrix}x_1\\0\end{bmatrix} + \begin{bmatrix}0\\x_2\end{bmatrix}\\ &= x_1\begin{bmatrix}1\\0\end{bmatrix} + x_2\begin{bmatrix}0\\1\end{bmatrix}\end{aligned}$$

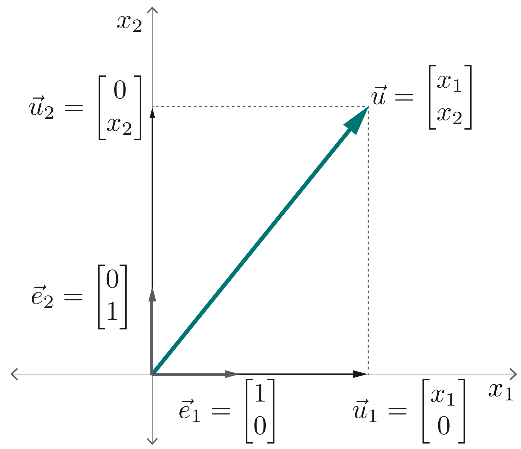
where $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$ are orthogonal to each other.

Define $\vec{e}_1 = \begin{bmatrix}1\\0\end{bmatrix}$ and $\vec{e}_2 = \begin{bmatrix}0\\1\end{bmatrix}$.

Define $\vec{u}_1 = \begin{bmatrix}x_1\\0\end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix}0\\x_2\end{bmatrix}$.

We can write the linear combination as follows:

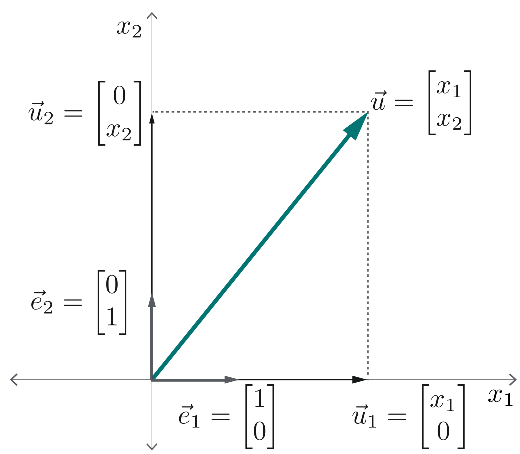
$$\begin{aligned}\vec{u} &= \vec{u}_1 + \vec{u}_2\\ &= x_1\vec{e}_1 + x_2\vec{e}_2\end{aligned}$$



Slide

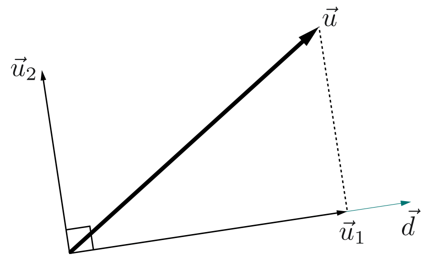
Intuition for Vector Projections Continued

Note that $\vec{u}_1$ lies along the vector $\vec{e}_1$ which is in the same direction as the x_1-axis.
Similarly, $\vec{u}_2$ lies along the vector $\vec{e}_2$ that is perpendicular to the vector $\vec{e}_1$.
Finally, note that $\vec{u} = \vec{u}_1 + \vec{u}_2$.
$\vec{u}_1$ is called a **projection** of $\vec{u}$ onto the vector $\vec{e}_1$.
$\vec{u}_2$ is called a **perpendicular** of $\vec{u}$ onto the vector $\vec{e}_1$.
To find the projection of $\vec{u}$ onto the vector $\vec{e}_1$, we simply drop the "shadow" of $\vec{u}$ onto the vector $\vec{e}_1$.



Slide

Projection onto a Vector



Suppose that we are given two vectors $\vec{u}$ and $\vec{d}$.
To find the **projection** of $\vec{u}$ onto $\vec{d}$, assume we have a light source at the tip of $\vec{u}$ and we drop the shadow of $\vec{u}$ onto $\vec{d}$.
We will call it $\vec{u}_1$.
Note that $\vec{u}_1$ is a scalar multiple of $\vec{d}$.
Next, we will find the **perpendicular** of $\vec{u}$ onto $\vec{d}$ by choosing a corresponding vector that is orthogonal to $\vec{d}$.
We will call this vector $\vec{u}_2$.
Notice that $\vec{u} = \vec{u}_1 + \vec{u}_2$.

In summary, given two vectors $\vec{u}, \vec{d} \in \mathbb{R}^n, \vec{d} \neq \vec{0}$, we can write $\vec{u} = \vec{u}_1 + \vec{u}_2$ where:

- $\vec{u}_1$is a scalar multiple of $\vec{d}$ and
- $\vec{u}_2$ is orthogonal to $\vec{d}$.

Definition of Projection

We have just seen the visual representation of $\vec{u} = \vec{u}_1 + \vec{u}_2$, where

- $\vec{u}_1$ is a scalar multiple of $\vec{d}$ and
- $\vec{u}_2$ is orthogonal to $\vec{d}$.

We now rearrange this equation to $\vec{u}_2 = \vec{u} - \vec{u}_1$. Since $\vec{u}_2$ is orthogonal to $\vec{d}$, we have

$$\vec{u}_2 \cdot \vec{d} = 0,$$

and since $\vec{u}_1$ is a scalar multiple of $\vec{d}$, we can write

$$\vec{u}_1 = t\vec{d}$$

for some $t \in \mathbb{R}$. Thus, given $\vec{u}_2 \cdot \vec{d} = 0$, if we can find t, then we know $\vec{u}_1$ and hence we will know $\vec{u}_2$. To find t, we do the following:

$$\begin{aligned}
0 &= \vec{u}_2 \cdot \vec{d} \\
&= (\vec{u} - \vec{u}_1) \cdot \vec{d} \\
&= \vec{u} \cdot \vec{d} - \vec{u}_1 \cdot \vec{d} \\
&= \vec{u} \cdot \vec{d} - (t\vec{d}) \cdot \vec{d}
\end{aligned}$$

Hence

$$\begin{aligned} 0 &= \vec{u}\cdot\vec{d} - t(\vec{d}\cdot\vec{d}) \\ &= \vec{u}\cdot\vec{d} - t||\vec{d}||^2 \end{aligned}$$

and since $\vec{d} \neq \vec{0}$, we can rearrange our equality as follows:

$$t = \frac{\vec{u}\cdot\vec{d}}{||\vec{d}||^2}$$

This leads us to the following definition:

Definition
Let $\vec{u}, \vec{d} \in \mathbb{R}^n, \vec{d} \neq \vec{0}$. The **projection** of $\vec{u}$ onto $\vec{d}$ is

$$\text{proj}_{\vec{d}}\vec{u} = \frac{\vec{u}\cdot\vec{d}}{||\vec{d}||^2}\vec{d}$$

and the projection of $\vec{u}$ perpendicular to $\vec{d}$ (or the **perpendicular** of $\vec{u}$ onto $\vec{d}$) is

$$\text{perp}_{\vec{d}}\vec{u} = \vec{u} - \text{proj}_{\vec{d}}\vec{u}$$

Note that for $\vec{u}, \vec{d} \in \mathbb{R}^n, \vec{d} \neq \vec{0}$,

$$\vec{u}_1 = \text{proj}_{\vec{d}}\vec{u} \quad \text{and} \quad \vec{u}_2 = \text{perp}_{\vec{d}}\vec{u}$$

Theorem 1
Let $\vec{u}, \vec{d} \in \mathbb{R}^n, \vec{d} \neq \vec{0}$. Then

1. $\text{proj}_{\vec{d}}\vec{u} = \dfrac{\vec{u}\cdot\vec{d}}{||\vec{d}||^2}\vec{d}$
2. $\text{perp}_{\vec{d}}\vec{u}$ is orthogonal to $\vec{d}$
3. $\text{proj}_{\vec{d}}\vec{u}$ is orthogonal to $\text{perp}_{\vec{d}}\vec{u}$

Proof

We have already verified (1) above.

For (2), we need to check that the dot product of $\text{perp}_{\vec{d}}\vec{u}$ and $\vec{d}$ is zero.

$$\begin{aligned} (\text{perp}_{\vec{d}}\cdot\vec{u})\cdot\vec{d} &= (\vec{u} - \text{proj}_{\vec{d}}\vec{u})\cdot\vec{d} \\ &= \vec{u}\cdot\vec{d} - \left(\frac{\vec{u}\cdot\vec{d}}{||\vec{d}||^2}\vec{d}\right)\cdot\vec{d} && \text{by (1)} \\ &= \vec{u}\cdot\vec{d} - \left(\frac{\vec{u}\cdot\vec{d}}{||\vec{d}||^2}\right)(\vec{d}\cdot\vec{d}) \\ &= \vec{u}\cdot\vec{d} - \vec{u}\cdot\vec{d} \\ &= 0 \end{aligned}$$

so $\text{perp}_{\vec{d}}\vec{u}$ is orthogonal to $\vec{d}$.

For (3), we need to check that the dot product of $\text{proj}_{\vec{d}}\vec{u}$ and $\text{perp}_{\vec{d}}\vec{u}$ is zero.

$$\begin{aligned}
\text{proj}_{\vec{d}}\vec{u} \cdot \text{perp}_{\vec{d}}\vec{u} &= \text{proj}_{\vec{d}}\vec{u} \cdot (\vec{u} - \text{proj}_{\vec{d}}\vec{u}) \\
&= \text{proj}_{\vec{d}}\vec{u} \cdot \vec{u} - \text{proj}_{\vec{d}}\vec{u} \cdot \text{proj}_{\vec{d}}\vec{u} \\
&= \left(\frac{\vec{u} \cdot \vec{d}}{||\vec{d}||^2} \vec{d} \right) \cdot \vec{u} - \left\| \text{proj}_{\vec{d}}\vec{u} \right\|^2 \\
&= \left(\frac{(\vec{u} \cdot \vec{d})^2}{||\vec{d}||^2} \right) - \left\| \text{proj}_{\vec{d}}\vec{u} \right\|^2 \\
&= \left\| \text{proj}_{\vec{d}}\vec{u} \right\|^2 - \left\| \text{proj}_{\vec{d}}\vec{u} \right\|^2 \\
&= 0
\end{aligned}$$

□

Making Connections

This applet displays vectors $\vec{u}$ and $\vec{d}$ along with the projection and perpendicular of $\vec{u}$ onto $\vec{d}$.

Instructions: Click and drag the vector $\vec{u}$ to observe how the projection and perpendicular change. Observe how statements 2 and 3 from Theorem 1 hold as you move the vector.

External resource: https://www.geogebra.org/material/iframe/id/v67htany/

Example

Let $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\vec{d} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$. Then,

$$\text{proj}_{\vec{d}}\vec{u} = \frac{\vec{u} \cdot \vec{d}}{||\vec{d}||^2}\vec{d} = \frac{-1+2+6}{1+1+4} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \frac{7}{6} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -7/6 \\ 7/6 \\ 7/3 \end{bmatrix}$$

$$\text{perp}_{\vec{d}}\vec{u} = \vec{u} - \text{proj}_{\vec{d}}\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -7/6 \\ 7/6 \\ 7/3 \end{bmatrix} = \begin{bmatrix} 13/6 \\ 5/6 \\ 2/3 \end{bmatrix}$$

Note that:

- $\text{proj}_{\vec{d}}\vec{u} = (7/6)\vec{d}$; the projection is a scalar multiple of $\vec{d}$
- $(\text{perp}_{\vec{d}}\vec{u}) \cdot \vec{d} = -13/6 + 5/6 + 4/3 = -8/6 + 8/6 = 0$; the perpendicular is orthogonal to $\vec{d}$
- $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -7/6 \\ 7/6 \\ 7/3 \end{bmatrix} + \begin{bmatrix} 13/6 \\ 5/6 \\ 2/3 \end{bmatrix} = \text{proj}_{\vec{d}}\vec{u} + \text{perp}_{\vec{d}}\vec{u}$

Exercise 1

A question appears in Mobius

Exercise 2

Let $\vec{u}, \vec{d} \in \mathbb{R}^n, \vec{d} \neq \vec{0}$. Prove that $\text{proj}_{\vec{d}}(\text{proj}_{\vec{d}}\vec{u}) = \text{proj}_{\vec{d}}\vec{u}$

A question appears in Mobius

Projections onto a Plane

Definition and Examples

We have already seen how to project a vector onto another vector. Here, we will see how to project a vector onto a plane in $\mathbb{R}^3$.

A slideshow appears in Mobius.

Slide

Projections onto a Plane

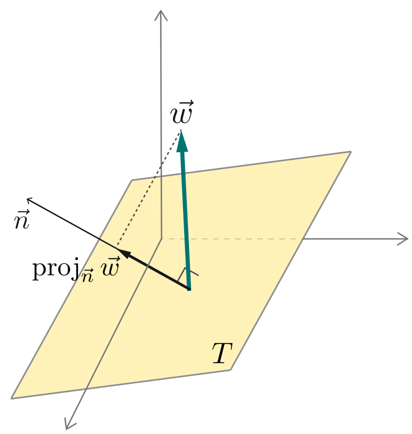
Let $\vec{u}$ and $\vec{v}$ be two non-parallel vectors in $\mathbb{R}^3$. Then the vector equation

$$\vec{x} = s\vec{u} + t\vec{v}, \quad s, t \in \mathbb{R}$$

defines a plane T passing through the origin.
Let $\vec{n} = \vec{u} \times \vec{v}$ be a normal vector to T.
Given an arbitrary vector $\vec{w}$ in $\mathbb{R}^3$, we can compute a projection of $\vec{w}$ onto $\vec{n}$ using the standard formula for the projection:

$$\operatorname{proj}_{\vec{n}} \vec{w} = \frac{\vec{w} \cdot \vec{n}}{\|\vec{n}\|^2}\vec{n}$$

By definition, this vector is parallel to $\vec{n}$ and hence orthogonal to the plane T.



Slide

Projections onto a Plane

We define the **perpendicular** of $\vec{w}$ onto T, denoted by $\operatorname{perp}_T \vec{w}$, as

$$\operatorname{perp}_T \vec{w} = \operatorname{proj}_{\vec{n}} \vec{w}$$

We define the **projection** of $\vec{w}$ onto T as

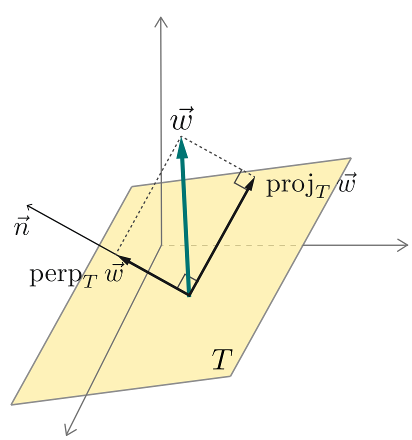
$$\operatorname{proj}_T \vec{w} = \vec{w} - \operatorname{perp}_T \vec{w}$$

Note that this definition implies that

$$\operatorname{proj}_T \vec{w} = \vec{w} - \operatorname{proj}_{\vec{n}} \vec{w} = \operatorname{perp}_{\vec{n}} \vec{w}$$

By construction, the vectors $\operatorname{proj}_T \vec{w}$ and $\operatorname{perp}_T \vec{w}$ are orthogonal and

$$\operatorname{proj}_T \vec{w} + \operatorname{perp}_T \vec{w} = \vec{w}$$



Making Connections

This applet displays a plane and a vector $\vec{u}$ along with the projection and perpendicular of $\vec{u}$ onto the plane.

Instructions: Click and drag the vector $\vec{u}$ to observe how the projection and perpendicular change.

External resource: https://www.geogebra.org/material/iframe/id/tebstckx/

Example 1

Let $\vec{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$. Then we can find a normal $\vec{n}$ to the plane containing $\vec{u}$ and $\vec{v}$ by computing the cross product:

$$\vec{n} = \vec{u} \times \vec{v} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \times \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ -2 \end{bmatrix}.$$

Therefore we can find the projection of $\vec{w}$ onto T as follows:

$$\begin{aligned}
\operatorname{proj}_T \vec{w} &= \operatorname{perp}_{\vec{n}} \vec{w} \\
&= \vec{w} - \operatorname{proj}_{\vec{n}} \vec{w} \\
&= \vec{w} - \frac{\vec{w} \cdot \vec{n}}{\|\vec{n}\|^2} \vec{n} \\
&= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \frac{5}{\left(\sqrt{30}\right)^2} \begin{bmatrix} 1 \\ 5 \\ -2 \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 1/6 \\ 5/6 \\ -1/3 \end{bmatrix} \\
&= \begin{bmatrix} 5/6 \\ 7/6 \\ 10/3 \end{bmatrix}
\end{aligned}$$

Remark
Note that in this course the projection of a vector onto a plane is defined only for planes that pass through the origin.

Example 2

Let $\vec{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, and let T be a plane with scalar equation $x_1 + 5x_2 - 2x_3 = 0$. We would like to compute $\operatorname{proj}_T \vec{w}$.

Note that since $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ satisfies the scalar equation, our plane passes through the origin, and so $\operatorname{proj}_T \vec{w}$ is well-defined.

We see that a normal $\vec{n}$ to the plane T is $\begin{bmatrix} 1 \\ 5 \\ -2 \end{bmatrix}$, so

$$\operatorname{proj}_T \vec{w} = \operatorname{perp}_{\vec{n}} \vec{w} = \begin{bmatrix} 5/6 \\ 7/6 \\ 10/3 \end{bmatrix}$$

as in the previous example.

Exercise

A question appears in Mobius

Unit 2

Systems of Linear Equations

2.1 - Systems of Linear Equations

Systems of Linear Equations

Examples of Systems of Linear Equations

In this section, we'll go over a few important definitions. When looking at examples, you may notice some familiar elements from your earlier work in mathematics.

Definition
A **linear equation** is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where

- $a_1, a_2, \ldots, a_n$ are real numbers;
- $x_1, x_2, \ldots, x_n$ are variables (or unknowns);
- b is the constant term.

We can group linear equations into a system of linear equations.

Definition
A **system of linear equations** is a collection of finitely many linear equations.

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

Solutions to Systems of Linear Equations

Let's first define the solution to a single linear equation as follows:

Definition

A **hyperplane** in $\mathbb{R}^n$is the set of all vectors $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ that satisfy a linear equation $a_1x_1 + \cdots + a_nx_n = b$, in the case when $a_1, \ldots, a_n$ are not all simultaneously equal to zero.

- In the case where $n = 1$, hyperplanes are called points.
- In the case where $n = 2$, hyperplanes are called lines.
- In the case where $n = 3$, hyperplanes are called planes.

Now, let's consider a system of linear equations. We are interested in the following questions:

- Does the system have a solution?
- If so, what are all the solutions?
- What do solutions to the system represent geometrically?

We will explore the answers to these questions in the following examples.

Example 1

Consider

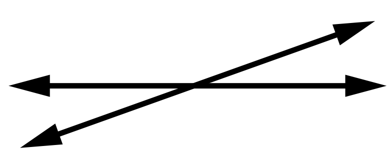
$$\begin{array}{rcrcl} ax_1 & + & bx_2 & = & c \\ dx_1 & + & ex_2 & = & f \end{array}$$

where $a, b, c, d, e, f \in \mathbb{R}$ such that a, b are not both zero and d, e are not both zero. Notice that the last two conditions mean that both equations in this system represent hyperplanes in $\mathbb{R}^2$; in $\mathbb{R}^2$, hyperplanes are simply lines.

Hence the solution set corresponds to the intersection of two lines. There are three possible options for what this set might look like:

- Parallel lines that do not intersect. In this case, there are no solutions.

- Nonparallel lines that intersect at a point. In this case, there is exactly one solution.



- Parallel lines that intersect. In this case, there are infinitely many solutions.

Given any system of linear equations, the same result will hold: we will either have no solutions, or exactly one solution, or infinitely many solutions. Geometrically, we will now have more options for what a set with infinitely many solutions might look like.

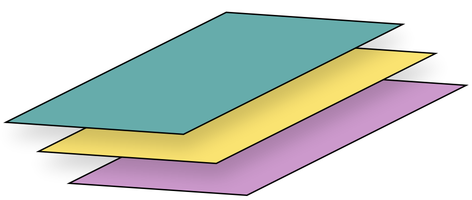
Example 2

Consider a system of three linear equations in three unknowns

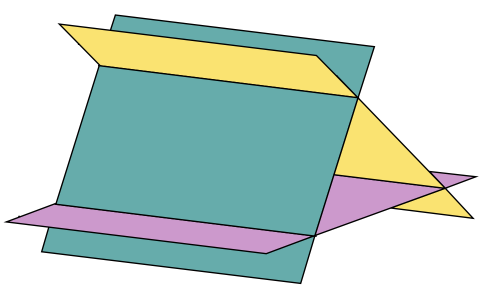
$$\begin{array}{ccccccc} a_1x_1 & + & a_2x_2 & + & a_3x_3 & = & d_1 \\ b_1x_1 & + & b_2x_2 & + & b_3x_3 & = & d_2 \\ c_1x_1 & + & c_2x_2 & + & c_3x_3 & = & d_3 \end{array}$$

where at least one of the coefficients in each equation is non-zero. In this case, every linear equation in this system represents a plane in $\mathbb{R}^3$. Hence the solution set corresponds to intersection of three planes. Here are some options for what this set might look like:

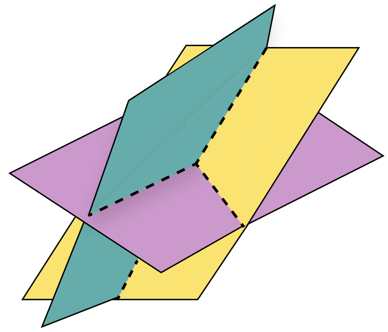
- At least two out of three planes are parallel to each other. In this case, there are no solutions.



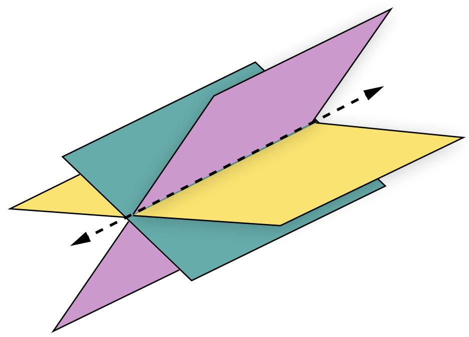
- Lines of intersection are parallel, but not coincident. In this case, there are no solutions.



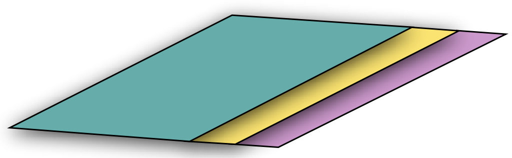
- Three planes intersect at a point. In this case, there is exactly one solution.



- Three planes intersect at a line. In this case, there are infinitely many solutions.



- Three planes are identical. In this case, there are infinitely many solutions that lie on a single plane.

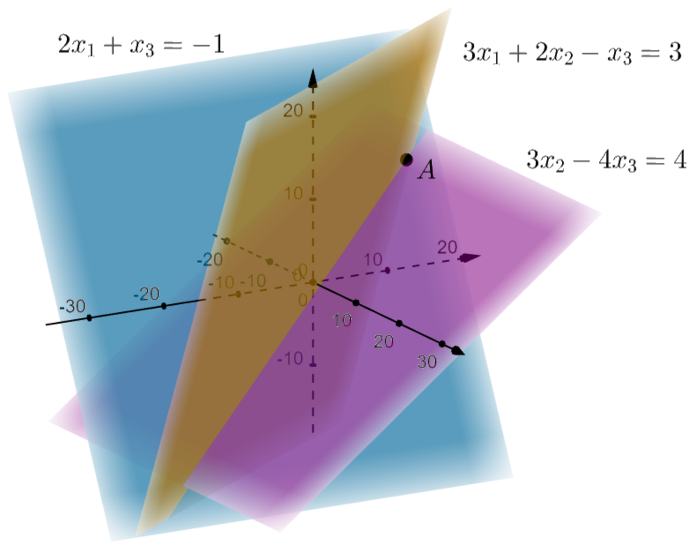


Example 3

Consider the following system of linear equations:

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

There is a hyperplane in $\mathbb{R}^3$ that satisfies each equation, so every solution $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ to this system of equations belongs to the intersection of three hyperplanes in $\mathbb{R}^3$. If we were to solve the system, we would see that it has the unique solution $\begin{bmatrix} -6 \\ 16 \\ 11 \end{bmatrix}$ (point A in the figure), which belongs to the intersection of three hyperplanes in $\mathbb{R}^3$. This would appear geometrically as follows:



Exercise 3

A question appears in Mobius

2.2 - Augmented and Coefficient Matrices, Reduced Row Echelon Form, and Gaussian Elimination

Augmented and Coefficient Matrices

Introduction

Our first introduction to matrices begins with this section, where we look at augmented and coefficient matrices associated with systems of linear equations. Note that we will study matrices in greater detail in the next unit.

A **matrix** is defined as a rectangular array of numbers.

One of the benefits of using matrices is that they allow us to represent systems more compactly.

Definition
Given a system of m linear equations in n unknowns,

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

the **augmented matrix** of the system is the $m \times (n+1)$ matrix with m rows and $n+1$ columns written as follows

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

The **coefficient matrix** of the system is the $m \times n$ matrix with m rows and n columns written as follows

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

We refer to $\vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$ as the **constant matrix** (or **constant vector**) of our system. If we denote the coefficient matrix above by A, then the augmented matrix is denoted by $\left[A \mid \vec{b}\right]$.

Let us see an example of augmented, coefficient and constant matrices associated with a system of linear equations.

Example 1

Write the augmented, coefficient, and constant matrices for the given system of linear equations:

$$\begin{array}{rcrcr} 4x_1 & - & 5x_2 & = & 7 \\ 2x_1 & + & x_2 & = & -3 \end{array}$$

Solution

- $\left[\begin{array}{cc|c} 4 & -5 & 7 \\ 2 & 1 & -3 \end{array}\right]$ is the **augmented matrix** of the system, denoted by $\left[A \mid \vec{b}\right]$.
- $\begin{bmatrix} 4 & -5 \\ 2 & 1 \end{bmatrix}$ is the **coefficient matrix** of the system, denoted by A.
- $\begin{bmatrix} 7 \\ -3 \end{bmatrix}$ is the **constant matrix** (or constant vector) of the system, denoted by $\vec{b}$.

Exercise

A question appears in Mobius

Example 2

Solve the following system of two linear equations in two unknowns:

$$\begin{array}{rcrcr} x_1 & + & 3x_2 & = & -1 \\ x_1 & + & x_2 & = & 3 \end{array}$$

Solution 1

First, eliminate the $x_1$in the second equation by subtracting the first equation from the second:

$$\begin{aligned} x_1 + 3x_2 &= -1 \\ x_1 + x_2 &= 3 \end{aligned} \quad \Rightarrow \quad \text{Subtract 1st equation from 2nd} \quad \Rightarrow \quad \begin{aligned} x_1 + 3x_2 &= -1 \\ -2x_2 &= 4 \end{aligned}$$

Next, isolate the x_2 in the second equation by multiplying the equation by $-1/2$:

$$\begin{aligned} x_1 + 3x_2 &= -1 \\ -2x_2 &= 4 \end{aligned} \quad \Rightarrow \quad \text{Multiply 2nd equation by -1/2} \quad \Rightarrow \quad \begin{aligned} x_1 + 3x_2 &= -1 \\ x_2 &= -2 \end{aligned}$$

Finally, eliminate the x_2 in the first equation by subtracting 3 times the second equation from the first:

$$\begin{aligned} x_1 + 3x_2 &= -1 \\ x_2 &= -2 \end{aligned} \quad \Rightarrow \quad \text{Subtract 3 times 2nd equation from 1st} \quad \Rightarrow \quad \begin{aligned} x_1 &= 5 \\ x_2 &= -2 \end{aligned}$$

From here, we conclude that the solution to the system is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

Solution 2

Now, let's solve this question again in a more compact way. We'll repeat the system here for convenience:

$$\begin{array}{rcrcr} x_1 & + & 3x_2 & = & -1 \\ x_1 & + & x_2 & = & 3 \end{array}$$

First, record coefficients and constant terms of our system in its augmented matrix

$$\left[\begin{array}{cc|c} 1 & 3 & -1 \\ 1 & 1 & 3 \end{array}\right]$$

We refer to the first and second rows of this matrix as R_1 and R_2, respectively.

Now apply the same operations as we did in Solution 1, but write the steps as follows:

$$\left[\begin{array}{cc|c} 1 & 3 & -1 \\ 1 & 1 & 3 \end{array}\right] \underset{R_2-R_1}{\sim} \left[\begin{array}{cc|c} 1 & 3 & -1 \\ 0 & -2 & 4 \end{array}\right] \underset{(-1/2)R_2}{\sim} \left[\begin{array}{cc|c} 1 & 3 & -1 \\ 0 & 1 & -2 \end{array}\right] \underset{R_1-3R_2}{\sim} \left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & -2 \end{array}\right]$$

Therefore we see again that the solution is:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

From the previous example, we see that by taking the augmented matrix of a linear system of equations, we can apply a sequence of operations that "reduce" it to an augmented matrix of a much simpler system, which has the same solution. In the next section we will learn about the algorithm that enables us to solve any system of linear equations.

(Reduced) Row Echelon Form

Elementary Row Operation

Definition
The term **elementary row operation** (**ERO**) corresponds to one of the three operations on the rows of a coefficient matrix or the augmented matrix of a system $\left[A \mid \vec{b}\right]$.
The three elementary row operations are:

- Swap the i-th and j-th rows, denoted $R_i \leftrightarrow R_j$,
- Add a scalar multiple of the i-th row to the j-th row, denoted $R_j + cR_i$, for $c \in \mathbb{R}$.
- Multiply the i-th row by a non-zero scalar $c \in \mathbb{R}$, denoted cR_i.

Elementary row operations have the important property that they preserve the solution set of the system. Given a second system, $\left[B \mid \vec{d}\right]$, we say that the systems $\left[A \mid \vec{b}\right]$ and $\left[B \mid \vec{d}\right]$ are **equivalent**, denoted by $\left[A \mid \vec{b}\right] \sim \left[B \mid \vec{d}\right]$, if they have the same solution set. Since elementary row operations preserve the solution set, any two matrices that can be transformed into each other via a sequence of elementary row operations will be equivalent.

In an earlier example, we used elementary row operations to reduce the complicated system

$$\left[\begin{array}{cc|c} 1 & 3 & -1 \\ 1 & 1 & 3 \end{array}\right]$$

to a simpler one

$$\left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & -2 \end{array}\right]$$

Thus

$$\left[\begin{array}{cc|c} 1 & 3 & -1 \\ 1 & 1 & 3 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & -2 \end{array}\right]$$

and the systems they represent

$$\begin{array}{rcrcr} x_1 & + & 3x_2 & = & -1 \\ x_1 & + & x_2 & = & 3 \end{array} \qquad \text{and} \qquad \begin{array}{rcrcr} x_1 & & & = & 5 \\ & & x_2 & = & -2 \end{array}$$

must have the same solution set. Clearly, the second system is easier to solve.

Exercise 1

A question appears in Mobius

Definition
We say that a system $\left[A \mid \vec{b}\right]$ is **consistent** if it has at least one solution, and we say that the system is **inconsistent** if it has no solutions.

Notice that the system

$$\begin{array}{rcrcr} x_1 & + & 3x_2 & = & -1 \\ x_1 & + & x_2 & = & 3 \end{array}$$

which we have been working with is consistent since it has exactly one solution $x_1 = 5$, $x_2 = -2$.

Example 1

Let's solve the following system of three linear equations in three unknowns

$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & 9x_3 & = & 31 \\ & & x_2 & + & 2x_3 & = & 8 \\ x_1 & & & + & 3x_3 & = & 10 \end{array}$$

Solution 1

To solve this system, we first write the augmented matrix and then perform elementary row operations to it

$$\left[\begin{array}{ccc|c} 2 & 1 & 9 & 31 \\ 0 & 1 & 2 & 8 \\ 1 & 0 & 3 & 10 \end{array}\right] \underset{R_1 \leftrightarrow R_3}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 2 & 1 & 9 & 31 \end{array}\right] \underset{R_3 - 2R_1}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & 1 & 3 & 11 \end{array}\right] \underset{R_3 - R_2}{\sim}$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right] \underset{R_1 - 3R_3}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right] \underset{R_2 - 2R_3}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

We conclude that

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Solution 2

Notice that we could have avoided the last two elementary row operations in Solution 1 and stopped at the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

which corresponds to the system

$$\begin{array}{ccccccc} x_1 & & & + & 3x_3 & = & 10 \\ & & x_2 & + & 2x_3 & = & 8 \\ & & & & x_3 & = & 3 \end{array}$$

Since we know the value of x_3, we can solve for x_2, and then for x_1:

$$x_2 = 8 - 2x_3 = 8 - 2(3) = 8 - 6 = 2$$

$$x_1 = 10 - 3x_3 = 10 - 3(3) = 10 - 9 = 1$$

to again arrive at $x_1 = 1$, $x_2 = 2$, $x_3 = 3$. This technique is called **back substitution**.

Definition of (Reduced) Row Echelon Form

Given a particular system of linear equations, we apply elementary row operations to transform it into another system, which is much easier to solve. We introduce two such types of simple systems. Before we proceed, note that we refer to the first non-zero entry in each row of a matrix as a **leading entry** (or a **pivot**).

Definition
A matrix is in **row echelon form** (**REF**) if

1. all rows that contain only zero entries (zero rows) are at the bottom, and
2. each leading entry is to the right of the leading entries above it.

A matrix is in **reduced row echelon form** (**RREF**) if it is in REF and

1. each leading entry is a 1 (called a **leading one**), and
2. each leading one is the only non-zero entry in its column.

Example 2

Here is an example of a matrix in row echelon form (REF), with leading entries bolded:

$$\begin{bmatrix} 3 & 2 & -1 & 3 \\ 0 & 0 & -1 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Here is an example of a matrix in reduced row echelon form (RREF), with leading entries bolded:

$$\begin{bmatrix} 1 & 2/3 & 0 & -2/3 \\ 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Remark

1. Note that if a matrix is in RREF, then it is in REF.
2. For any matrix we have many REFs, but the RREF is unique.

Exercise 2

A question appears in Mobius

It turns out that it is always possible to convert a given system into REF or RREF. The process of calculation of REF of a matrix is called **row reduction** or **Gaussian elimination.** When row reducing the augmented matrix of a linear system of equations

- we aim first for REF, then
- once we have a matrix in REF, we may either use back substitution, or continue using elementary row operations until we are in RREF, then
- we can simply read off the solution.

Looking back at our previous example, we note that

$$\left[\begin{array}{ccc|c} 2 & 1 & 9 & 31 \\ 0 & 1 & 2 & 8 \\ 1 & 0 & 4 & 10 \end{array}\right] \sim \underbrace{\left[\begin{array}{ccc|c} 1 & 0 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right]}_{\text{REF}} \sim \underbrace{\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]}_{\text{REF and RREF}}$$

Exercise 3

A question appears in Mobius

Example 3

Let's solve the following system of three linear equations in three unknowns

$$\begin{array}{rcrcrcr} 3x_1 & + & x_2 & & & = & 10 \\ 2x_1 & + & x_2 & + & x_3 & = & 6 \\ -3x_1 & + & 4x_2 & + & 15x_3 & = & -20 \end{array}$$

Solution

To solve this system, we perform elementary row operations on the augmented matrix:

$$\left[\begin{array}{ccc|c} 3 & 1 & 0 & 10 \\ 2 & 1 & 1 & 6 \\ -3 & 4 & 15 & -20 \end{array}\right] \underset{R_1-R_2}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 4 \\ 2 & 1 & 1 & 6 \\ -3 & 4 & 15 & -20 \end{array}\right] \underset{\substack{R_2-2R_1 \\ R_3+3R_1}}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 4 \\ 0 & 1 & 3 & -2 \\ 0 & 4 & 12 & -8 \end{array}\right] \underset{R_3-4R_2}{\sim}$$

$$\underbrace{\left[\begin{array}{ccc|c} 1 & 0 & -1 & 4 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right]}_{\text{RREF}}$$

The resulting system is

$$\begin{array}{rcrcrcr} x_1 & - & & & x_3 & = & 4 \\ & & x_2 & + & 3x_3 & = & -2 \\ & & & & 0 & = & 0 \end{array}$$

Note that the third equation is always true.

Given an REF of an augmented matrix we define the following:

- The variables corresponding to a column with a leading entry are called **leading variables**.
  In this example, the leading variables are x_1, x_2 since the first two columns of the last matrix have leading entries in them.

- The variables corresponding to columns without a leading entry are called **free variables** (or non-leading variables).
  In this example, the variable x_3 is free.

We assign **parameters** to the free variables and **solve** for the leading variables in terms of these parameters as follows:

$$\begin{aligned} x_1 &= 4 + x_3 \\ x_2 &= -2 - 3x_3 \qquad t \in \mathbb{R} \\ x_3 &= t \end{aligned}$$

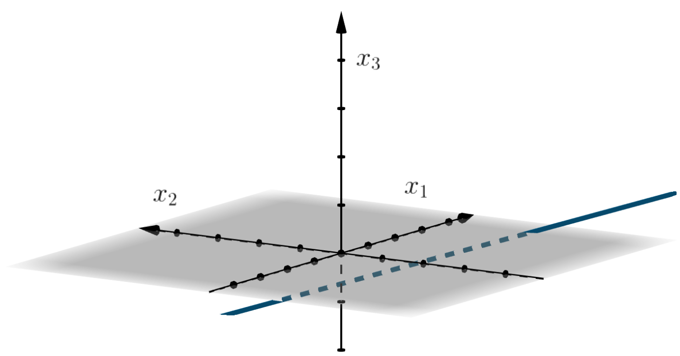
so

$$\begin{aligned} x_1 &= 4 + t \\ x_2 &= -2 - 3t \qquad t \in \mathbb{R} \\ x_3 &= t \end{aligned}$$

or equivalently

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

The existence of a parameter in the solution means that we have **infinitely many solutions** (one for each value of the parameter). Also, notice that the solution is a line in $\mathbb{R}^3$:



Example 4

Let us solve the following system of two linear equations in four unknowns:

$$\begin{array}{ccccccccc} x_1 & + & 6x_2 & & & - & x_4 & = & -1 \\ & & & & x_3 & + & 2x_4 & = & 7 \end{array}$$

Solution

We see that the augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 6 & 0 & -1 & -1 \\ 0 & 0 & 1 & 2 & 7 \end{array}\right]$$

is already in RREF. We see that x_1 and x_3 are leading variables, as the first and third columns have leading entries in them. We also see that x_2 and x_4 are free variables and will each be assigned **different** parameters.

The solution set for this system has the form

$$\begin{aligned} x_1 &= -1 - 6s + t \\ x_2 &= s \\ x_3 &= 7 - 2t \\ x_4 &= t \end{aligned} \qquad s, t \in \mathbb{R}$$

or equivalently

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 7 \\ 0 \end{bmatrix} + s \begin{bmatrix} -6 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix}, \qquad s, t \in \mathbb{R}$$

We recognize this as a vector equation with two parameters. This corresponds to a plane in $\mathbb{R}^4$.

Example 5

Let's solve the following system of three linear equations in three unknowns:

$$\begin{array}{rcrcrcr} 2x_1 & + & 12x_2 & - & 8x_3 & = & -4 \\ 2x_1 & + & 13x_2 & - & 6x_3 & = & -5 \\ -2x_1 & - & 14x_2 & + & 4x_3 & = & 7 \end{array}$$

Solution

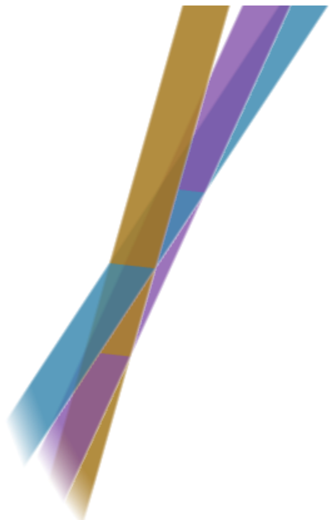
To solve this system, we perform elementary row operations on the augmented matrix:

$$\left[\begin{array}{ccc|c} 2 & 12 & -8 & -4 \\ 2 & 13 & -6 & -5 \\ -2 & -14 & 4 & 7 \end{array}\right] \underset{\substack{R_2 - R_1 \\ R_3 + R_1}}{\sim} \left[\begin{array}{ccc|c} 2 & 12 & -8 & -4 \\ 0 & 1 & 2 & -1 \\ 0 & -2 & -4 & 3 \end{array}\right] \underset{R_3 + 2R_2}{\sim} \left[\begin{array}{ccc|c} 2 & 12 & -8 & -4 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

The resulting system is

$$\begin{array}{rcrcrcr} 2x_1 & + & 12x_2 & - & 8x_3 & = & -4 \\ & & x_2 & + & 2x_3 & = & -1 \\ & & & & 0 & = & 1 \end{array}$$

Note that the last equation $0 = 1$ is never true. This means that our system has **no solutions**, i.e., it is **inconsistent**. Here, our equations are planes in $\mathbb{R}^3$ without a common intersection:



If the process of row reduction of an augmented matrix reveals a row of the form

$$\left[\begin{array}{ccc|c} 0 & \cdots & 0 & c \end{array}\right]$$

with $c \neq 0$, then the system is **inconsistent**. In this case, there is no need to continue applying elementary row operations. Note that for $c \neq 0$ such a row indicates that there is a leading entry in the last column. In other words, if the last column of an augmented matrix contains a leading entry, then the system has to be inconsistent.

Gaussian Elimination

REF and RREF Algorithms

As mentioned earlier, the algorithms for obtaining an REF or RREF for a given matrix is called **Gaussian elimination**, or **row reduction**. From earlier examples, you may have already observed some patterns in what operations are chosen and when. Below are the algorithms.

REF Algorithm
To produce an REF of a given matrix:

1. Find the first non-zero column in a given matrix. Perform an elementary row operation to obtain a **leading entry** at the top of the column.
2. Make all the entries **below** the leading entry equal to zero using elementary row operations.
3. Repeat steps 1 and 2 for the submatrix to the right of the current column and below the row with the leading entry.

RREF Algorithm
To produce an RREF of a given matrix:

1. Find the first non-zero column in a given matrix. Perform an elementary row operation to obtain a **leading one** at the top of the column.

2. Make all the entries **above and below** the leading one equal to zero using elementary row operations.

3. Repeat steps 1 and 2 for the submatrix to the right of the current column and below the row with the leading one.

Example 1

Let us find an REF of the matrix

$$\begin{bmatrix} 0 & 0 & 3 & 2 \\ 0 & 2 & -1 & 3 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

using the algorithm described above.

Solution

We notice that the second column is the first column that contains non-zero entries. Following the algorithm, we need to obtain a leading entry at the top of the column. This can be achieved either by performing $R_1 \leftrightarrow R_2$ or by performing $R_1 \leftrightarrow R_3$. We choose the latter

$$\begin{bmatrix} 0 & 0 & 3 & 2 \\ 0 & 2 & -1 & 3 \\ 0 & 1 & 1 & 0 \end{bmatrix} \underset{R_1 \leftrightarrow R_3}{\sim} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix}$$

Next, we have to make the entries below our leading entry equal to zero. This can be achieved by subtracting R_1 twice from R_2:

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix} \underset{R_2 - 2R_1}{\sim} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix}$$

Now we turn our attention to the submatrix, $\begin{bmatrix} -3 & 3 \\ 3 & 2 \end{bmatrix}$, which is to the right of the second column and below the first row:

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix}$$

We see that the first non-zero column already has a leading entry. Thus the remaining task is to make all entries below it equal to zero. This can be done by performing $R_3 + R_2$:

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix} \underset{R_3 + R_2}{\sim} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

Since the matrix we got is in REF, our algorithm terminates.

Remark
Row reduction is a common source of arithmetic mistakes, so be very cautious when doing it, especially when applying multiple ERO's in one step. One can do this only in the following two cases:

- No row is modified more than once in the same step.

For example, it is not allowed to perform $2R_1 + 3R_3$ in the same step since R_1 is being multiplied by 2 **and** $3R_3$ is added to $2R_1$ in the same step.

- No row is both modified and used to modify another row in the same step.

For example, it is not allowed to perform $R_1 + R_2$ and $R_3 - R_1$ in the same step.
Notice that using one row to modify other rows is allowed. For example, we can use $R_2 - R_1$ and $R_3 + 4R_1$ in one step.

Example 2

Let us find an RREF of the matrix

$$\begin{bmatrix} 0 & 0 & 3 & 2 \\ 0 & 2 & -1 & 3 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

using the algorithm described above.

Solution

We notice that the second column is the first column that contains non-zero entries. Following the algorithm, we need to obtain a leading one at the top of the column. This can be achieved either by performing $R_1 \leftrightarrow R_2$ or by performing $R_1 \leftrightarrow R_3$. We choose the latter:

$$\begin{bmatrix} 0 & 0 & 3 & 2 \\ 0 & 2 & -1 & 3 \\ 0 & 1 & 1 & 0 \end{bmatrix} \underset{R_1 \leftrightarrow R_3}{\sim} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix}$$

Next, we have to make the entries below our leading one equal to zero. This can be achieved by subtracting R_1 twice from R_2:

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix} \underset{R_2 - 2R_1}{\sim} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix}$$

Now we turn our attention to the submatrix, $\begin{bmatrix} -3 & 3 \\ 3 & 2 \end{bmatrix}$, which is to the right of the second column and below the first row:

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & \color{red}{-3} & \color{red}{3} \\ 0 & 0 & \color{red}{3} & \color{red}{2} \end{bmatrix}$$

We see that the first non-zero column has a non-zero entry at the top, which we have to turn into a leading one:

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix} \underset{(-1/3)R_2}{\sim} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 3 & 2 \end{bmatrix}$$

It remains to make all entries **above and below** the leading one equal to zero. This can be done by performing $R_3 - 3R_2$ and $R_1 - R_2$:

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 3 & 2 \end{bmatrix} \underset{R_3-3R_2}{\sim} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 5 \end{bmatrix} \underset{R_1-R_2}{\sim} \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

Next, we turn our attention to the submatrix, $\begin{bmatrix} 5 \end{bmatrix}$, which is to the right of the third column and below the second row:

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

We see that its only column has a non-zero entry, which we have to turn into a leading one

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 5 \end{bmatrix} \underset{(1/5)R_3}{\sim} \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Finally, we make all of the entries above the leading one equal to zero

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underset{R_1-R_3}{\sim} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underset{R_2+R_3}{\sim} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In the end, we obtain the matrix in RREF.

Linear Combinations of Vectors

Given arbitrary vectors $\vec{v}, \vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$ it is natural to ask whether $\vec{v}$ can be written as a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$. This problem naturally reduces to the problem of solving a system of n linear equations in k unknowns, which in turn can be solved using Gaussian elimination.

Example 3

Let

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} -6 \\ -12 \end{bmatrix}, \quad \vec{u} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

Let us see whether it is possible to express:

1. $\vec{u}$ as a linear combination of $\vec{x}$ and $\vec{y}$, and/or
2. $\vec{v}$ as a linear combination of $\vec{x}$ and $\vec{y}$.

Solution

For $\vec{u}$, we seek $c_1, c_2 \in \mathbb{R}$ so that $\vec{u} = c_1\vec{x} + c_2\vec{y}$. That is,

$$\begin{bmatrix} 4 \\ 8 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -6 \\ -12 \end{bmatrix}$$

This leads to the linear system

$$\begin{array}{rcrcl} c_1 & - & 6c_2 & = & 4 \\ 2c_1 & - & 12c_2 & = & 8 \end{array}$$

which has augmented matrix

$$\left[\begin{array}{cc|c} 1 & -6 & 4 \\ 2 & -12 & 8 \end{array}\right]$$

Now for $\vec{v}$ we seek $d_1, d_2 \in \mathbb{R}$ so that $\vec{v} = d_1\vec{x} + d_2\vec{y}$. That is,

$$\begin{array}{rcrcl} d_1 & - & 6d_2 & = & 4 \\ 2d_1 & - & 12d_2 & = & 5 \end{array}$$

which has augmented matrix

$$\left[\begin{array}{cc|c} 1 & -6 & 4 \\ 2 & -12 & 5 \end{array}\right]$$

As the coefficient matrices are the same for both systems, we may solve them together

$$\left[\begin{array}{cc|cc} 1 & -6 & 4 & 4 \\ 2 & -12 & 8 & 5 \end{array}\right] \underset{R_2-2R_1}{\sim} \left[\begin{array}{cc|cc} 1 & -6 & 4 & 4 \\ 0 & 0 & 0 & -3 \end{array}\right]$$

For $\vec{u}$, we have

$$\left[\begin{array}{cc|c} 1 & -6 & 4 \\ 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{array}{l} c_1 = 4 + 6t \\ c_2 = t \end{array} \quad t \in \mathbb{R}$$

This means that, yes, we can express $\vec{u}$ as a linear combination of $\vec{x}$ and $\vec{y}$, and there are infinitely many ways to do so. For example,

$$\begin{array}{ll} t = 0 \Rightarrow & \vec{u} = 4\vec{x} + 0\vec{y} \\ t = 1 \Rightarrow & \vec{u} = 10\vec{x} + \vec{y} \end{array}$$

etc.

For $\vec{v}$, we have

$$\left[\begin{array}{cc|c} 1 & -6 & 4 \\ 2 & -12 & 5 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -6 & 4 \\ 0 & 0 & -3 \end{array}\right]$$

which is inconsistent as seen by the second row.

Hence $\vec{v}$ cannot be expressed as a linear combination of $\vec{x}$ and $\vec{y}$.

Remark
Depending on the problem that you are solving, sometimes it is more efficient to use REF in place of RREF. For example, if a question asks whether a vector is a linear combination of a set of vectors or if your goal is to determine whether a particular system is consistent, it is easier to obtain REF, as it enables us to answer such questions easily.
If a question asks to determine all the solutions of a system, then depending on the problem RREF might be a better choice.

A question appears in Mobius

A question appears in Mobius

Rank and Its Properties

Introduction

Definition
The **rank** of a matrix A, denoted $\operatorname{rank}(A)$, is the number of leading entries in any REF of A.

Let $\left[A \mid \vec{b}\right]$ be a system of m linear equations in n variables. The ranks of A and $\left[A \mid \vec{b}\right]$ allow us to deduce information about the solvability of $\left[A \mid \vec{b}\right]$. More precisely,

1. $\operatorname{rank}(A) < \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right)$ if and only if the system has no solutions;
2. $\operatorname{rank}(A) = \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right) = n$ if and only if the system has exactly one solution;
3. $\operatorname{rank}(A) = \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right) < n$ if and only if the system has infinitely many solutions; and
4. $\operatorname{rank}(A) = m$ if and only if the system $[A \mid \vec{c}]$ has a solution for every $\vec{c}$, where $\vec{c}$ is an arbitrary vector in $\mathbb{R}^m$.

Let's have a closer look at what these properties really mean.

A slideshow appears in Mobius.

Slide

Rank and Its Properties

Let A be an $m \times n$ matrix.
We define the **rank** of A as the number of leading entries in any row echelon form of A.
Let $\vec{b}$ be a vector in $\mathbb{R}^m$ and consider the system of linear equations with the augmented matrix $\left[A \mid \vec{b}\right]$.
Then $\operatorname{rank}(A)$ and $\operatorname{rank}\left(\left[A \mid \vec{b}\right]\right)$ provide us with important information about the number of solutions of this system.

Slide

Property 1

$$\operatorname{rank}(A) < \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right) \text{ if and only if the system } \left[A \mid \vec{b}\right] \text{ has no solutions}$$

Example: Consider the system of $m = 4$ linear equations in $n = 3$ unknowns

$$\begin{array}{rcrcrcl} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ x_1 & + & 3x_2 & + & x_3 & = & 0 \\ 3x_1 & - & x_2 & + & 3x_3 & = & 0 \\ & - & 5x_2 & & & = & 1 \end{array}$$

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 3 & 1 \\ 3 & -1 & 3 \\ 0 & -5 & 0 \end{bmatrix} \qquad \left[A \mid \vec{b}\right] = \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 1 & 3 & 1 & 0 \\ 3 & -1 & 3 & 0 \\ 0 & -5 & 0 & 1 \end{array}\right]$$

$$\operatorname{RREF}(A) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \operatorname{RREF}\left(\left[A \mid \vec{b}\right]\right) = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore $\operatorname{rank}(A) = 2 < 3 = \operatorname{rank}([A \mid \vec{b}])$, which means that the system is inconsistent (has no solutions).
Alternative: The system $\left[A \mid \vec{b}\right]$ has no solutions if and only if the RREF of the augmented matrix contains a row where all of the entries are zero, except the last one.

Slide

Property 1 Continued

$$\text{rank}(A) < \text{rank}\left(\left[A \mid \vec{b}\right]\right) \text{ if and only if the system } \left[A \mid \vec{b}\right] \text{ has no solutions}$$

Example: Consider the system of $m = 4$ linear equations in $n = 3$ unknowns

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 2 \\ x_1 & + & 3x_2 & + & x_3 & = & -3 \\ 3x_1 & - & x_2 & + & 3x_3 & = & 1 \\ & - & 5x_2 & & & = & 5 \end{array}$$

$$\text{RREF}(A) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{RREF}\left(\left[A \mid \vec{b}\right]\right) = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore $\text{rank}(A) = 2 = \text{rank}\left(\left[A \mid \vec{b}\right]\right)$, which means that the system has at least one solution.

Note: It is not necessary to row reduce A and $\left[A \mid \vec{b}\right]$ separately. By row reducing $\left[A \mid \vec{b}\right]$ we automatically row reduce A.

Slide

Property 2

$$\text{rank}(A) = \text{rank}\left(\left[A \mid \vec{b}\right]\right) = n \text{ if and only if the system } \left[A \mid \vec{b}\right] \text{ has exactly one solution}$$

Example: Consider the system of $m = 4$ linear equations in $n = 3$ unknowns

$$\begin{array}{rcrcrcr} 2x_1 & & & - & x_3 & = & 1 \\ x_1 & + & 2x_2 & + & 2x_3 & = & -3 \\ 2x_1 & - & 5x_2 & + & 4x_3 & = & 3 \\ 3x_1 & + & 2x_2 & + & x_3 & = & -2 \end{array}$$

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 2 & 2 \\ 2 & -5 & 4 \\ 3 & 2 & 1 \end{bmatrix} \qquad \left[A \mid \vec{b}\right] = \left[\begin{array}{ccc|c} 2 & 0 & -1 & 1 \\ 1 & 2 & 2 & -3 \\ 2 & -5 & 4 & 3 \\ 3 & 2 & 1 & -2 \end{array}\right]$$

$$\text{RREF}(A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{RREF}\left(\left[A \mid \vec{b}\right]\right) = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1/5 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -3/5 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Since $\text{rank}(A) = \text{rank}\left(\left[A \mid \vec{b}\right]\right) = 3 = n$, the system has exactly one solution.

From $\text{RREF}\left(\left[A \mid \vec{b}\right]\right)$ we see that this solution is

$$x_1 = 1/5, \quad x_2 = -1, \quad x_3 = -3/5$$

Slide

Property 3

$$\operatorname{rank}(A) = \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right) < n \text{ if and only if the system } \left[A \mid \vec{b}\right] \text{ has infinitely many solutions}$$

Example
Consider the system of $m = 4$ linear equations in $n = 3$ unknowns

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 2 \\ x_1 & + & 3x_2 & + & x_3 & = & -3 \\ 3x_1 & - & x_2 & + & 3x_3 & = & 1 \\ & - & 5x_2 & & & = & 5 \end{array}$$

$$\operatorname{RREF}(A) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \operatorname{RREF}\left(\left[A \mid \vec{b}\right]\right) = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Since $\operatorname{rank}(A) = \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right) = 2 < 3 = n$, there are infinitely many solutions.
From $\operatorname{RREF}\left(\left[A \mid \vec{b}\right]\right)$ we see that the solution set is

$$x_1 = -x_3, \quad x_2 = -1, \quad x_3 \in \mathbb{R}$$

Slide

Property 4

$$\operatorname{rank}(A) = m \text{ if and only if the system } [A \mid \vec{c}] \text{ has a solution for any } \vec{c} \in \mathbb{R}^m$$

Example
Consider the system of $m = 3$ linear equations in $n = 3$ unknowns

$$\begin{array}{rcrcrcr} 2x_1 & & & - & x_3 & = & c_1 \\ x_1 & + & 2x_2 & + & 2x_3 & = & c_2 \\ 2x_1 & - & 5x_2 & + & 4x_3 & = & c_3 \end{array}$$

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 2 & 2 \\ 2 & -5 & 4 \end{bmatrix}$$

$$\operatorname{RREF}(A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Since $\operatorname{rank}(A) = 3 = m$, the system has a solution for any $c_1, c_2, c_3 \in \mathbb{R}$.

Slide

Property 4 Continued

$$\text{rank}(A) = m \text{ if and only if the system } [A \mid \vec{c}] \text{ has a solution for any } \vec{c} \in \mathbb{R}^m$$

Example
Consider the system of $m = 4$ linear equations in $n = 3$ unknowns

$$\begin{array}{rcrcrcl} x_1 & - & 2x_2 & + & x_3 & = & c_1 \\ x_1 & + & 3x_2 & + & x_3 & = & c_2 \\ 3x_1 & - & x_2 & + & 3x_3 & = & c_3 \\ & - & 5x_2 & & & = & c_4 \end{array}$$

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 3 & 1 \\ 3 & -1 & 3 \\ 0 & -5 & 0 \end{bmatrix}$$

$$\text{RREF}(A) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Since $\text{rank}(A) = 2 < 4 = m$, there exist c_1, c_2, c_3, c_4 such that the system has no solutions.
For example, we may take

$$c_1 = 0, \quad c_2 = 0, \quad c_3 = 0, \quad c_4 = 1$$

Example 1

Consider the following system, $\left[A \mid \vec{b}\right]$, of three linear equations in three variables:

$$\left[A \mid \vec{b}\right] = \left[\begin{array}{ccc|c} 2 & 1 & 9 & 31 \\ 0 & 1 & 2 & 8 \\ 1 & 0 & 3 & 10 \end{array}\right]$$

Determine whether the system has no solutions, one solution, or infinitely many solutions.

Solution

Applying row reduction, we get

$$\left[\begin{array}{ccc|c} 2 & 1 & 9 & 31 \\ 0 & 1 & 2 & 8 \\ 1 & 0 & 3 & 10 \end{array}\right] \underset{R_3 \leftrightarrow R_1}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 2 & 1 & 9 & 31 \end{array}\right] \underset{R_3 - 2R_1}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & 1 & 3 & 11 \end{array}\right] \underset{R_3 - R_2}{\sim} \underbrace{\left[\begin{array}{ccc|c} 1 & 0 & 3 & 10 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right]}_{\text{REF}}$$

The REF of $\left[A \mid \vec{b}\right]$ has three leading entries, so $\text{rank}\left(\left[A \mid \vec{b}\right]\right) = 3$.

The coefficient matrix of the REF, A, also has three leading entries, so $\text{rank}(A) = 3$.

Since $\text{rank}(A) = \text{rank}\left(\left[A \mid \vec{b}\right]\right) = 3$, then, from the properties listed above, we deduce that the system of equations is consistent. Moreover, this system has to have exactly one solution, because $\text{rank}(A) = n = 3$.

This solution turns out to be $x_1 = 1, x_2 = 2, x_3 = 3$.

Example 2

Consider the coefficient matrix

$$B = \begin{bmatrix} 2 & 0 & 1 & 3 & 4 \\ 5 & 1 & 6 & -7 & 3 \end{bmatrix}$$

Determine whether the system has no solutions, one solution, or infinitely many solutions.

Solution

Applying row reduction, we get

$$\begin{bmatrix} 2 & 0 & 1 & 3 & 4 \\ 5 & 1 & 6 & -7 & 3 \end{bmatrix} \underset{R_2-2R_1}{\sim} \begin{bmatrix} 2 & 0 & 1 & 3 & 4 \\ 1 & 1 & 4 & -13 & -5 \end{bmatrix} \underset{R_1 \leftrightarrow R_2}{\sim} \begin{bmatrix} 1 & 1 & 4 & -13 & -5 \\ 2 & 0 & 1 & 3 & 4 \end{bmatrix} \underset{R_2-2R_1}{\sim}$$

$$\underbrace{\begin{bmatrix} 1 & 1 & 4 & -13 & -5 \\ 0 & -2 & -7 & 29 & 14 \end{bmatrix}}_{\text{REF}}$$

Since the REF above has two leading entries, we conclude that $\text{rank}(B) = 2$.

We see that the system $[B \mid \vec{c}]$will remain consistent no matter which vector $\vec{c} \in \mathbb{R}^2$ we choose, because $\text{rank}(B) = 2 = m$. We can also see that each such system will have infinitely many solutions.

Example 3

Consider the matrix

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$$

Determine whether the system has no solutions, one solution, or infinitely many solutions.

Solution

Applying row reduction, we get

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \underset{R_2-2R_1}{\sim} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}}_{\text{REF}}$$

Since the REF above has one leading entry, we conclude that $\text{rank}(C) = 1$.

This examples shows the importance of reducing to REF before counting the leading entries: the coefficient matrix

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$$

has two leading entries, but $\text{rank}(C) = 1$.

Notice that not every system of the form $[C \mid \vec{c}]$ will be consistent, because $\text{rank}(C) = 1 \neq 2 = m$. For instance, the system $\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 2 & 4 & 6 & 0 \end{array}\right]$ is inconsistent.

Despite that, those systems that **are** consistent, such as the system $\left[C \mid \vec{0}\right]$, will have infinitely many solutions because $\operatorname{rank}(C) = \operatorname{rank}\left(\left[C \mid \vec{0}\right]\right) = 1 < 3 = n$.

Example 4

For which values of the parameters $k, \ell \in \mathbb{R}$ does the system

$$\begin{array}{ccccc} 2x_1 & + & 6x_2 & = & 5 \\ 4x_1 & + & (k+15)x_2 & = & \ell+8 \end{array}$$

have:

- no solutions,
- a unique solution, and
- infinitely many solutions.

Solution

Let

$$A = \begin{bmatrix} 2 & 6 \\ 4 & k+15 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 5 \\ \ell+8 \end{bmatrix}$$

We carry $\left[A \mid \vec{b}\right]$ to REF (or what appears to be REF):

$$\left[\begin{array}{cc|c} 2 & 6 & 5 \\ 4 & k+15 & \ell+8 \end{array}\right] \underset{R_2-2R_1}{\sim} \left[\begin{array}{cc|c} 2 & 6 & 5 \\ 0 & k+3 & \ell-2 \end{array}\right]$$

If $k+3 \neq 0$, then $\operatorname{rank}(A) = 2 = \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right)$, so the system is consistent. Furthermore, since $\operatorname{rank}(A) = 2 = n$, we see that this solution has to be unique.

If $k+3 = 0$, then our last augmented matrix takes the form

$$\left[\begin{array}{cc|c} 2 & 6 & 5 \\ 0 & 0 & \ell-2 \end{array}\right]$$

Now,

- If $\ell - 2 \neq 0$, then $\operatorname{rank}(A) = 1 \neq 2 = \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right)$, so the system is inconsistent; that is, it has no solutions.
- If $\ell - 2 = 0$, then $\operatorname{rank}(A) = 1 = \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right)$, so the system is consistent. Since $\operatorname{rank}(A) = \operatorname{rank}\left(\left[A \mid \vec{b}\right]\right) = 1 < 2 = n$, we see that such a system has infinitely many solutions.

In summary, our system has:

- no solutions when $k = -3$ and $\ell \neq 2$,
- a unique solution when $k \neq -3$, and

- infinitely many solutions when $k = -3$ and $\ell = 2$.

A question appears in Mobius

A question appears in Mobius

2.3 - Homogeneous and Non-Homogeneous Systems

Homogeneous and Non-Homogeneous Systems

Homogeneous Systems

Definitions
A linear equation is called **homogeneous** if it is of the form

$$a_1x_1 + \cdots + a_nx_n = 0$$

where $a_1, \ldots, a_n \in \mathbb{R}$.
A system of linear equations $\left[A \mid \vec{0}\right]$ is called **homogeneous** if all of its linear equations are homogeneous.

Example 1

Which of the following two linear equations is homogeneous?

$$(1) \quad 2x_1 - 3x_2 + 5x_3 = 1$$

or

$$(2) \quad 4x_1 + 6x_2 - x_3 = 0$$

Solution

The second equation is homogeneous because it is of the form $a_1x_1 + \cdots + a_nx_n = 0$.

Example 2

Which of the following two systems of linear equations is homogeneous?

$$(1) \quad \begin{array}{rcrcrcr} -5x_1 & + & 2x_2 & - & 3x_3 & = & 0 \\ & & x_2 & - & 13x_3 & = & 0 \end{array}$$

or

$$(2) \quad \begin{array}{rcrcrcr} 8x_1 & + & x_2 & - & 8x_3 & = & 1 \\ 25x_1 & + & 2x_2 & - & 17x_3 & = & 0 \end{array}$$

Solution

The first system is homogeneous, since **all** of its equations are homogeneous.

Example 3

Solve the following homogeneous system of linear equations:

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 0 \\ & & 3x_2 & - & x_3 & = & 0 \end{array}$$

Solution

Applying row reduction, we get

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & 3 & -1 & 0 \end{array}\right] \underset{(1/3)R_2}{\sim} \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & 1 & -1/3 & 0 \end{array}\right] \underset{R_1 - R_2}{\sim} \left[\begin{array}{rrr|r} 1 & 0 & 4/3 & 0 \\ 0 & 1 & -1/3 & 0 \end{array}\right]$$

so

$$\begin{array}{ll} x_1 = (-4/3)t & \\ x_2 = (1/3)t & \qquad t \in \mathbb{R} \\ x_3 = t & \end{array}$$

or equivalently

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = t \begin{bmatrix} -4/3 \\ 1/3 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

Note that

- Taking $t = 0$ gives the trivial solution $x_1 = x_2 = x_3 = 0$.
- Fractions can be eliminated as follows:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = t \begin{bmatrix} -4/3 \\ 1/3 \\ 1 \end{bmatrix} = \frac{t}{3} \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix} = s \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix}, \quad s \in \mathbb{R}$$

where $s = t/3$. We let the parameter s "absorb" the factor of $1/3$. This is not necessary, but it is useful if one wishes to eliminate fractions.

Remark
Note that, during the row reduction process, the very last column of an augmented matrix $\left[A \mid \vec{0}\right]$ remains unchanged. Consequently, we can omit the column of zeros, and row reduce only the coefficient matrix A instead of the augmented matrix.

Example 4

Solve the following homogeneous system of linear equations:

$$\begin{array}{rcrcrcrcl} 4x_1 & - & 2x_2 & + & 3x_3 & + & 5x_4 & = & 0 \\ 8x_1 & - & 4x_2 & + & 6x_3 & + & 11x_4 & = & 0 \\ -4x_1 & + & 2x_2 & - & 3x_3 & - & 7x_4 & = & 0 \end{array}$$

Solution

In view of the above remark, we apply row reduction to the coefficient matrix instead of the augmented matrix:

$$\begin{bmatrix} 4 & -2 & 3 & 5 \\ 8 & -4 & 6 & 11 \\ -4 & 2 & -3 & -7 \end{bmatrix} \underset{\substack{R_2-2R_1 \\ R_3+R_1}}{\sim} \begin{bmatrix} 4 & -2 & 3 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix} \underset{R_3+2R_2}{\sim} \begin{bmatrix} 4 & -2 & 3 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \underset{R_1-5R_2}{\sim}$$

$$\begin{bmatrix} 4 & -2 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \underset{(1/4)R_1}{\sim} \begin{bmatrix} 1 & -1/2 & 3/4 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

so,

$$\begin{aligned} x_1 &= \tfrac{1}{2}s - \tfrac{3}{4}t, \\ x_2 &= s, \\ x_3 &= t, \\ x_4 &= 0, \end{aligned} \qquad s, t \in \mathbb{R}$$

or equivalently

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = s \begin{bmatrix} 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3/4 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad s, t \in \mathbb{R}$$

Noting that $s = t = 0$ gives the trivial solution $x_1 = x_2 = x_3 = x_4 = 0$, we see that the solution set includes the origin. Since there are two parameters s and t associated with two linearly independent vectors, we can see that the solution set is a plane through the origin in $\mathbb{R}^4$.

Exercise 1

A question appears in Mobius

Non-Homogeneous Systems

Definitions

A linear equation is called **non-homogeneous** if it is of the form

$$a_1x_1 + \cdots + a_nx_n = c$$

where $a_1, \ldots, a_n, c \in \mathbb{R}$ and $c \neq 0$.

A system of linear equations $\left[A \mid \vec{b}\right]$ is called **non-homogeneous** if at least one of its equations is non-homogeneous, or equivalently $\vec{b} \neq \vec{0}$.

The system $\left[A \mid \vec{0}\right]$ is called the **associated homogeneous system** of $\left[A \mid \vec{b}\right]$.

The following example demonstrates an interesting connection between the solution set of a non-homogeneous system $\left[A \mid \vec{b}\right]$ and the solution set of its associated homogeneous system $\left[A \mid \vec{0}\right]$.

Example 5

Let us solve the following non-homogeneous system of linear equations:

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 1 \\ & & 3x_2 & - & x_3 & = & 3 \end{array}$$

Applying row reduction, we get

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 3 & -1 & 3 \end{array}\right] \underset{(1/3)R_2}{\sim} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1/3 & 1 \end{array}\right] \underset{R_1-R_2}{\sim} \left[\begin{array}{ccc|c} 1 & 0 & 4/3 & 0 \\ 0 & 1 & -1/3 & 1 \end{array}\right]$$

so

$$\begin{array}{l} x_1 = (-4/3)t \\ x_2 = 1 + (1/3)t \ , \quad t \in \mathbb{R} \\ x_3 = t \end{array}$$

or equivalently

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -4/3 \\ 1/3 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

Note that this system has the same coefficient matrix as the homogeneous system from Example 3. In fact, the system from Example 3 is precisely the associated homogeneous system of this system.

Non-Homogeneous System

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 1 \\ & & 3x_2 & - & x_3 & = & 3 \end{array}$$

Solution set

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\text{particular solution}} + t \underbrace{\begin{bmatrix} -4/3 \\ 1/3 \\ 1 \end{bmatrix}}_{\text{complete solution}}$$

Associated Homogeneous System

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 0 \\ & & 3x_2 & - & x_3 & = & 0 \end{array}$$

Solution set

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = t \begin{bmatrix} -4/3 \\ 1/3 \\ 1 \end{bmatrix}$$

As stated in the following theorem, such a relationship between the solution sets of the two systems holds in general.

Theorem 1
Let $\left[A \mid \vec{b}\right]$ be a non-homogeneous system of linear equations. If

- we know one solution to a non-homogeneous system, which we call a **particular solution**, and
- we know the complete solution to the corresponding homogeneous system,

then the complete solution to the non-homogeneous system is the particular solution plus the complete solution to the associated homogeneous system.

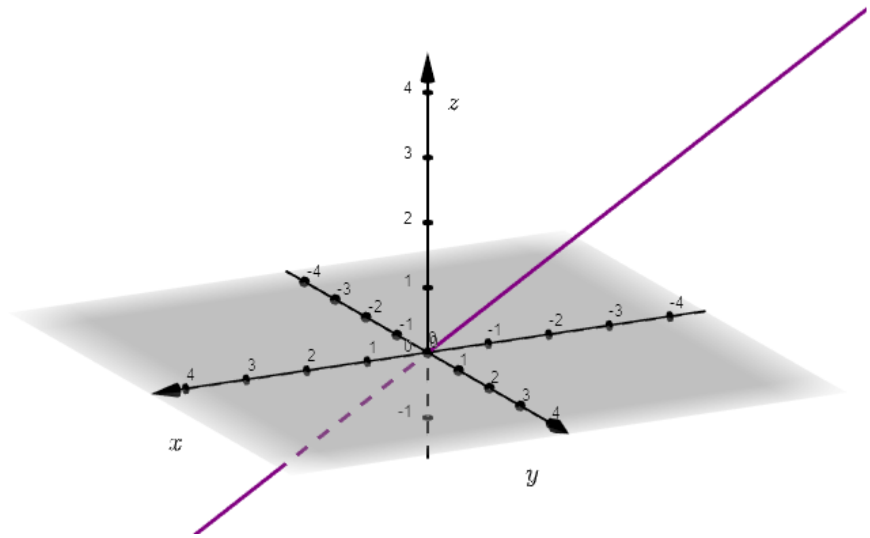
Consider the homogeneous system

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 3 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 4/3 & 0 \\ 0 & 1 & -1/3 & 0 \end{array}\right]$$

which has solution set

$$t \begin{bmatrix} -4/3 \\ 1/3 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

Geometrically, this solution set forms a line in $\mathbb{R}^3$ which passes through the origin.



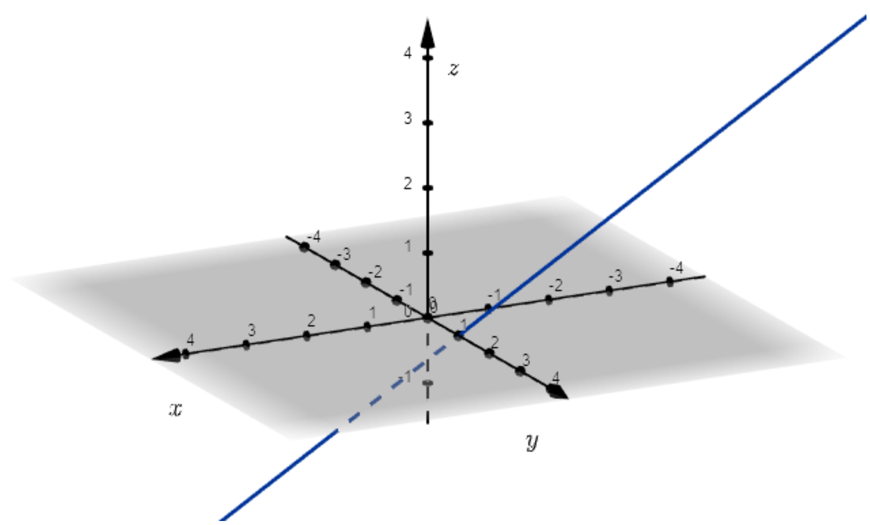
Now, consider the associated non-homogeneous system

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 3 & -1 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 4/3 & 0 \\ 0 & 1 & -1/3 & 1 \end{array}\right]$$

which has solution set

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -4/3 \\ 1/3 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

Geometrically, this solution set forms a line in $\mathbb{R}^3$. Notice that this line does not pass through the origin, but intersects the y-axis at the point $(0, 1, 0)$.



Making Connections

In this activity, you can directly compare the solution sets obtained above, as well as a few others.

To display one or more solution sets, select them. To remove a solution set, select it again.

You can select and drag the image to rotate it.

External resource: https://www.geogebra.org/material/iframe/id/dtkwywt6/

Notice that:

- The solution sets form parallel lines.
- The solution set for the homogeneous system passes through the origin, while the solution set for the non-homogeneous system is the same line, but shifted or translated by the value of the particular solution. In this example, the particular solution was $(0, 1, 0)$, so the line corresponding to the non-homogeneous system is shifted one unit in the direction of the positive y-axis.
- The solution sets of different non-homogeneous systems are lines which are parallel to the solution set of the homogeneous system, with an appropriate shift in the x, y, and z directions.

Example 6

If $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is a solution to a non-homogeneous system, and $t\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ is the complete solution to the associated homogeneous system, find 2 more solutions to the non-homogeneous system.

Solution

We know that:

- $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is one solution to the non-homogeneous system and
- $t\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $t \in \mathbb{R}$ is the complete solution to the associated homogeneous system.

It follows from the theorem stated above that the complete solution to the non-homogeneous system is the sum of these two solutions and has general form

$$\underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}_{\text{particular solution}} + t \underbrace{\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}}_{\text{complete solution}}, \quad t \in \mathbb{R}$$

Since we can choose any real value for t, we can choose $t = 1$ and $t = -1$ to get two solutions:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}$$

and

$$\begin{bmatrix}1\\2\\3\end{bmatrix} - \begin{bmatrix}4\\5\\6\end{bmatrix} = \begin{bmatrix}-3\\-3\\-3\end{bmatrix}$$

Example 7

If $\begin{bmatrix}1\\2\\3\end{bmatrix}$ and $\begin{bmatrix}4\\5\\6\end{bmatrix}$ are solutions to the same non-homogeneous system, find one solution to the associated homogeneous system.

Solution

Suppose that A is the coefficient matrix associated to the non-homogeneous system. Since $\begin{bmatrix}1\\2\\3\end{bmatrix}$ and $\begin{bmatrix}4\\5\\6\end{bmatrix}$ are both solutions, they satisfy $A\begin{bmatrix}1\\2\\3\end{bmatrix} = \vec{b}$ and $A\begin{bmatrix}4\\5\\6\end{bmatrix} = \vec{b}$ for some vector $b \in \mathbb{R}^3$. Subtracting the equations, we have

$$A\begin{bmatrix}1\\2\\3\end{bmatrix} - A\begin{bmatrix}4\\5\\6\end{bmatrix} = \vec{0}$$

which we can simplify to

$$A\left(\begin{bmatrix}1\\2\\3\end{bmatrix} - \begin{bmatrix}4\\5\\6\end{bmatrix}\right) = \vec{0}$$

Hence the vector $\begin{bmatrix}1\\2\\3\end{bmatrix} - \begin{bmatrix}4\\5\\6\end{bmatrix} = \begin{bmatrix}-3\\-3\\-3\end{bmatrix}$ is a solution to the associated homogeneous system since $A\begin{bmatrix}-3\\-3\\-3\end{bmatrix} = \vec{0}$.

Exercise 2

A question appears in Mobius

Exercise 3

A question appears in Mobius

Unit 3

Matrix Algebra

3.1 - Matrix Algebra

Matrix Algebra

Introduction

In an earlier lesson, you were introduced to augmented, coefficient, and constant matrices as a way to represent and work with linear systems. In this lesson, we will have a closer look at what matrices are, their properties, and operations that can be performed on them.

Definition
A **matrix** is defined as a rectangular array of numbers.

Example 1

$A = \begin{bmatrix} 8 & -2 \\ a & 3 \\ 23 & -5 \end{bmatrix}$ is a matrix. Since A has **3 rows** and **2 columns**, A is a 3×2 matrix.

$B = \begin{bmatrix} 0 & 0 \\ 0 & \sin\pi \end{bmatrix}$ is a matrix. Since B has 2 rows and 2 columns, B is a 2×2 matrix, also called a square matrix.

We write $A_{m\times n}$ if we want to emphasize that A has m rows and n columns. For the examples just provided, here's how this would look:

$$A_{3\times 2} = \begin{bmatrix} 8 & -2 \\ a & 3 \\ 23 & -5 \end{bmatrix}$$

$$B_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & \sin\pi \end{bmatrix}$$

Let's now look at a more detailed definition of a matrix.

Definition
A **matrix** A is a rectangular array of numbers with m rows and n columns.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

For $1 \leq i \leq m$ and $1 \leq j \leq n$, a_{ij} is the entry of A lying in the ith row and jth column of A and is referred to as the (i, j)-entry of A. Sometimes, we refer to the coefficient a_{ij} by writing $(A)_{ij}$.

Example 2

Let $A = \begin{bmatrix} 3 & \pi & 7 & 4 & 0 & 1 \\ 0 & -7 & 9 & -2 & 3 & \sqrt{5} \\ 1/2 & 17 & -3 & \sqrt{11} & -1 & 0 \\ 1 & 0 & 0 & 0 & \pi/2 & 8 \end{bmatrix}$.

Then

- 7 is the entry of A lying in the 1st row and 3rd column, also written as $(A)_{13}$;
- -2 is the entry of A lying in the 2nd row and 4th column, also written as $(A)_{24}$;
- 8 is the entry of A lying in the 4th row and 6th column, also written as $(A)_{46}$; and
- -3 is the entry of A lying in the 3rd row and 3rd column, also written at $(A)_{33}$.

Exercise 1

A question appears in Mobius

We say that $m \times n$ matrices A and B are equal if $(A)_{ij} = (B)_{ij}$ for every $1 \leq i \leq m$ and $1 \leq j \leq n$. In other words, matrices A and B are equal if their entries are equal.

Next, we introduce the definitions of the zero matrix and the identity matrix.

Definition
The $m \times n$ matrix with all zero entries is called a **zero matrix**, denoted by $0_{m\times n}$ (or just 0 if the size is clear).

Definition
The $n \times n$ **identity matrix**, denoted by I_n (or just I if the size is clear), is the square matrix with 1s on the main diagonal and 0s elsewhere.

Example 3

The following are examples of identity matrices:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Addition

Matrices are said to be compatible when they have the same size, i.e., the same number of rows and the same number of columns. Compatible matrices can be added (or subtracted).

Definition
For matrices A and B we define their **sum** as $(A+B)_{ij} = (A)_{ij} + (B)_{ij}$. That is, we add corresponding entries.

Example 4

The sum of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ is

$$\begin{aligned} (A+B) &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \\ &= \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} \\ &= \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix} \end{aligned}$$

Exercise 2

A question appears in Mobius

Scalar Multiplication

In addition to summing matrices, we can also multiply a matrix by a scalar.

Definition
Let A be a matrix and let $k \in \mathbb{R}$. We define **scalar multiplication** as

$$(kA)_{ij} = k(A)_{ij}$$

In particular, $1A$ is denoted by A and $(-1)A$ is denoted by $-A$.
Notice that for any A and any $k \in \mathbb{R}$, $0A = 0_{m\times n}$ and $k\,0_{m\times n} = 0_{m\times n}$.

Example 5

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Then

$$\begin{aligned} 3A &= 3\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 3\cdot 1 & 3\cdot 2 \\ 3\cdot 3 & 3\cdot 4 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix} \end{aligned}$$

Exercise 3

A question appears in Mobius

As in the following example, we can combine addition and scalar multiplication of matrices:

Example 6

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. Then,

$$\begin{aligned} 3A - 5B &= 3\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 5\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix} - \begin{bmatrix} 25 & 30 \\ 35 & 40 \end{bmatrix} \\ &= \begin{bmatrix} -22 & -24 \\ -26 & -28 \end{bmatrix} \end{aligned}$$

Exercise 4

A question appears in Mobius

The following theorem summarizes some useful properties of matrices.

Theorem 1: Properties of Matrices
Let A, B, C be $m \times n$ matrices and k, $l \in \mathbb{R}$. Then

1. $A + B$ is an $m \times n$ matrix
2. $A + B = B + A$
3. $A + (B + C) = (A + B) + C$
4. There is a matrix, denoted $0_{m\times n}$, such that $A + 0_{m\times n} = 0_{m\times n} + A = A$
5. For every A, there is a matrix $(-A)$, such that $A + (-A) = (-A) + A = 0_{m\times n}$
6. kA is an $m \times n$ matrix
7. $(k\,l)A = k(lA)$
8. $(k + l)A = kA + lA$
9. $k(A + B) = kA + kB$
10. $1A = A$

3.2 - Special Matrices

Transpose Matrix

Definition of the Transpose Matrix

In this lesson, we'll discuss some special types of matrices, starting with the transpose matrix.

Definition
Let A be an $m \times n$ matrix. The **transpose** of A, denoted A^T, is an $n \times m$ matrix such that $(A^T)_{ij} = (A)_{ji}$. This means that the rows of A^T are the columns of A, while the columns of A^T are the rows of A.

In the examples that follow, notice how the rows and columns of a matrix and its transpose are switched.

Example 1

Find the transpose of $A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Solution

$A^T = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

Example 2

Find the transpose of $B = \begin{bmatrix} 1 & 4 & 8 \end{bmatrix}$.

Solution

$$B^T = \begin{bmatrix} 1 \\ 4 \\ 8 \end{bmatrix}$$

Example 3

Find the transpose of $C = \begin{bmatrix} 1 & 4 \\ -1 & 3 \end{bmatrix}$.

Solution

$$C^T = \begin{bmatrix} 1 & -1 \\ 4 & 3 \end{bmatrix}$$

Exercise

A question appears in Mobius

Properties of the Transpose Matrix

The transpose of a matrix has the following properties.

Proposition: Properties of the Transpose Matrix
Let A, B be $m \times n$ matrices and $k \in \mathbb{R}$. Then,

1. A^T is an $n \times m$ matrix
2. $(A^T)^T = A$
3. $(A + B)^T = A^T + B^T$
4. $(kA)^T = kA^T$

We can observe Property 1 in the previous examples. Let's look at some examples of the remaining properties.

Example 4

Given $A = \begin{bmatrix} 6 & 2 & 5 \\ -4 & -1 & 7 \end{bmatrix}$, find $(A^T)^T$.

Solution

By Property 2,

$$(A^T)^T = A = \begin{bmatrix} 6 & 2 & 5 \\ -4 & -1 & 7 \end{bmatrix}$$

We could also have concluded this by first taking the transpose of A, and then taking the transpose of our result.

Example 5

Determine $\left(2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right)^T$.

Solution

Using property 4, we have

$$\left(2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right)^T = 2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^T = 2\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 4 & 8 \end{bmatrix}$$

Example 6

Solve for A if

$$\left(2A^T - 3\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}\right)^T = \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}$$

Solution

Evaluating the transpose that is outside of the brackets gives

$$\begin{aligned}
(2A^T)^T - \left(3\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}\right)^T &= \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix} && \text{by property 3} \\
2(A^T)^T - 3\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}^T &= \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix} && \text{by property 4} \\
2A - 3\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}^T &= \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix} && \text{by property 2} \\
2A - 3\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} &= \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix} && \text{by definition of transpose} \\
2A &= \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix} + 3\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \\
2A &= \begin{bmatrix} 5 & 0 \\ 5 & 5 \end{bmatrix} \\
A &= \begin{bmatrix} \frac{5}{2} & 0 \\ \frac{5}{2} & \frac{5}{2} \end{bmatrix}
\end{aligned}$$

Symmetric Matrices

Definition of a Symmetric Matrix

Let's now look at another special type of matrix, a symmetric matrix.

Definition
A matrix A is **symmetric** if $A^T = A$.

If A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix. So $A^T = A$ implies $m = n$. Thus a symmetric matrix must be square.

Example 1

Determine if the matrix $A = \begin{bmatrix} 1 & 6 \\ 6 & 9 \end{bmatrix}$ is symmetric.

Solution

$A^T = \begin{bmatrix} 1 & 6 \\ 6 & 9 \end{bmatrix} = A$, so A is symmetric.

Example 2

Determine if matrix $B = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 4 & 5 \\ 3 & 6 & 7 \end{bmatrix}$ is symmetric.

Solution

$B^T = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 4 & 6 \\ 3 & 5 & 7 \end{bmatrix} \neq B$, so B is not symmetric.

Above we mentioned that every symmetric matrix must be square. However, as this example shows, the converse of this statement is not true: B is a square matrix, but it is not symmetric.

Exercise

A question appears in Mobius

Example 3

Prove that if the matrices A and B are symmetric, then $kA + lB$ is symmetric for any $k, l \in \mathbb{R}$.

Solution

We must show that $(kA + lB)^T = kA + lB$. We have

$$\begin{aligned} (kA + lB)^T &= (kA)^T + (lB)^T && \text{by property 3} \\ &= kA^T + lB^T && \text{by property 4} \\ &= kA + lB && \text{since A and B are symmetric} \end{aligned}$$

Hence $kA + lB = (kA + lB)^T$ and $kA + lB$ is symmetric.

Triangular Matrices

Definition of a Triangular Matrix

Another special type of square matrices are triangular matrices.

Definition
An $n \times n$ matrix A is **lower triangular** if the (i, j)-entry of A is zero whenever $i < j$.

Example 1

The matrices $A = \begin{bmatrix} 1 & 0 \\ 7 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 6 & 0 \\ 1 & 1 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 0 & 0 & 0 \\ 5 & 0 & 0 \\ 2 & -2 & 0 \end{bmatrix}$ are lower triangular.

We can envison this as follows:

$$A = \begin{bmatrix} \color{red}{1} & 0 \\ \color{red}{7} & \color{red}{1} \end{bmatrix}, B = \begin{bmatrix} \color{red}{1} & 0 & 0 \\ \color{red}{-1} & \color{red}{6} & 0 \\ \color{red}{1} & \color{red}{1} & \color{red}{1} \end{bmatrix}, \text{ and } C = \begin{bmatrix} \color{red}{0} & 0 & 0 \\ \color{red}{5} & \color{red}{0} & 0 \\ \color{red}{2} & \color{red}{-2} & \color{red}{0} \end{bmatrix}.$$

Analogously, we define upper triangular matrices:

Definition
An $n \times n$ matrix A is **upper triangular** if the (i, j)-entry of A is zero whenever $i > j$.

Example 2

The matrices $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} -5 & 9 & \pi \\ 0 & -3 & 1 \\ 0 & 0 & 8 \end{bmatrix}$, and $C = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$ are upper triangular.

We can envision this as follows:

$$A = \begin{bmatrix} 0 & \color{red}{1} \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} \color{red}{-5} & \color{red}{9} & \color{red}{\pi} \\ 0 & \color{red}{-3} & \color{red}{1} \\ 0 & 0 & \color{red}{8} \end{bmatrix}, \text{ and } C = \begin{bmatrix} \color{red}{0} & \color{red}{1} & \color{red}{1} \\ 0 & \color{red}{0} & \color{red}{-1} \\ 0 & 0 & \color{red}{0} \end{bmatrix}.$$

Exercise

A question appears in Mobius

3.3 - Matrix-Vector Multiplication

Matrix-Vector Multiplication

Introduction

We can view a **vector** $\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \in \mathbb{R}^m$ as an $m \times 1$ **matrix**.

We can also view an $m \times n$ **matrix** A as a collection of n **vectors** in $\mathbb{R}^m$.

If we let $\vec{a}_1, \ldots, \vec{a}_n \in \mathbb{R}^m$ denote the columns of A, then we can write $A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}$. This indicates that:

- The 1st column of A is represented by the vector $\vec{a}_1 \in \mathbb{R}^m$.
- The 2nd column of A is represented by the vector $\vec{a}_2 \in \mathbb{R}^m$.
- $\vdots$
- The nth column of A is represented by the vector $\vec{a}_n \in \mathbb{R}^m$.

We now define matrix-vector multiplication $A\vec{x}$ as follows.

Definition

If $A = \begin{bmatrix} \vec{a}_1 & \ldots & \vec{a}_n \end{bmatrix}$ is an $m \times n$ matrix, such that $\vec{a}_1, \ldots, \vec{a}_n \in \mathbb{R}^m$and $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$, then
matrix-vector multiplication is defined by $A\,\vec{x} = x_1\,\vec{a}_1 + \cdots + x_n\,\vec{a}_n$.
Notice that the result of matrix-vector multiplication is a vector $A\vec{x} \in \mathbb{R}^m$.

Let us see an example of a matrix-vector multiplication.

Example 1

Multiply $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ by $\vec{x} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$.

Solution

$$\begin{aligned} A\vec{x} &= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \\ &= 1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + (-1) \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 6 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 4 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \end{bmatrix} + \begin{bmatrix} 6 \\ 12 \end{bmatrix} \\ &= \begin{bmatrix} 5 \\ 11 \end{bmatrix} \end{aligned}$$

This result is a vector in $\mathbb{R}^2$.

Example 2

Compute $A\vec{x}$, where

$$A = \begin{bmatrix} 1 & -1 & 6 \\ 0 & 2 & 1 \\ 4 & -3 & 2 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Solution

We have

$$A\vec{x} = 1\begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} + 1\begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} + 2\begin{bmatrix} 6 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1(1) + 1(-1) + 2(6) \\ 1(0) + 1(2) + 2(1) \\ 1(4) + 1(-3) + 2(2) \end{bmatrix} = \begin{bmatrix} 12 \\ 4 \\ 5 \end{bmatrix}$$

Note that in both of the above examples, the number of rows and columns in both A and $\vec{x}$ determines whether or not they can be multiplied together. The rule for this is covered in the next lesson.

Exercise 1

A question appears in Mobius

Properties of Matrix Multiplication

The properties of matrix-vector multiplication are summarized in the following theorem.

Theorem 1: Properties of Matrix-Vector Multiplication
Let A and B be $m \times n$ matrices, let $\vec{x}$, $\vec{y} \in \mathbb{R}^n$, and let $k \in \mathbb{R}$. Then

1. $A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y}$
2. $A(k\vec{x}) = k(A\vec{x})$
3. $(A+B)\vec{x} = A\vec{x} + B\vec{x}$

We'll look at each of these properties in the examples below.

Example 4

Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 1 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Verify that $A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y}$.

Solution

On one side, we have

$$\begin{aligned} A(\vec{x}+\vec{y}) &= \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right) \\ &= \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} \\ &= 3 \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} + (-1) \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 10 \\ -1 \end{bmatrix} \end{aligned}$$

On the other side,

$$\begin{aligned} A\vec{x} + A\vec{y} &= \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} \\ &= \left(1 \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} + (-1) \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right) + \left(2 \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right) \\ &= \begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 6 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 10 \\ -1 \end{bmatrix} \end{aligned}$$

Therefore, $A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y}$.

Example 5

Let

$$A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & 1 & 0 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}, \quad k = -2.$$

Verify that $A(k\vec{x}) = k(A\vec{x})$.

Solution

On one hand, we have

$$\begin{aligned} A(k\vec{x}) &= \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & 1 & 0 \end{bmatrix} \left(-2 \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix} \right) \\ &= \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -2 \\ 6 \\ -2 \end{bmatrix} \\ &= (-2) \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} + 6 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (-2) \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 6 \\ 6 \end{bmatrix} \end{aligned}$$

On the other hand,

$$\begin{aligned} k(A\vec{x}) &= (-2) \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix} \\ &= (-2) \left(1 \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} + (-3) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix} \right) \\ &= (-2) \begin{bmatrix} -4 \\ -3 \\ -3 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 6 \\ 6 \end{bmatrix} \end{aligned}$$

Therefore, $A(k\vec{x}) = k(A\vec{x})$.

Example 6

Let

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$$

Verify that $(A + B)\vec{x} = A\vec{x} + B\vec{x}$.

Solution

On one hand, we have

$$\begin{aligned}(A+B)\vec{x} &= \left(\begin{bmatrix}1 & -2 & 1\\ 1 & 1 & -1\end{bmatrix} + \begin{bmatrix}1 & 0 & 1\\ -2 & -1 & 0\end{bmatrix}\right)\begin{bmatrix}1\\ -2\\ 3\end{bmatrix}\\ &= \begin{bmatrix}2 & -2 & 2\\ -1 & 0 & -1\end{bmatrix}\begin{bmatrix}1\\ -2\\ 3\end{bmatrix}\\ &= 1\begin{bmatrix}2\\ -1\end{bmatrix} + (-2)\begin{bmatrix}-2\\ 0\end{bmatrix} + 3\begin{bmatrix}2\\ -1\end{bmatrix}\\ &= \begin{bmatrix}12\\ -4\end{bmatrix}\end{aligned}$$

On the other hand,

$$\begin{aligned}A\vec{x}+B\vec{x} &= \begin{bmatrix}1 & -2 & 1\\ 1 & 1 & -1\end{bmatrix}\begin{bmatrix}1\\ -2\\ 3\end{bmatrix} + \begin{bmatrix}1 & 0 & 1\\ -2 & -1 & 0\end{bmatrix}\begin{bmatrix}1\\ -2\\ 3\end{bmatrix}\\ &= \left(1\begin{bmatrix}1\\ 1\end{bmatrix} + (-2)\begin{bmatrix}-2\\ 1\end{bmatrix} + 3\begin{bmatrix}1\\ -1\end{bmatrix}\right) + \left(1\begin{bmatrix}1\\ -2\end{bmatrix} + (-2)\begin{bmatrix}0\\ -1\end{bmatrix} + 3\begin{bmatrix}1\\ 0\end{bmatrix}\right)\\ &= \begin{bmatrix}8\\ -4\end{bmatrix} + \begin{bmatrix}4\\ 0\end{bmatrix}\\ &= \begin{bmatrix}12\\ -4\end{bmatrix}\end{aligned}$$

Therefore, $(A+B)\vec{x} = A\vec{x}+B\vec{x}$.

As the following proposition suggests, a matrix-vector multiplication can be used to "extract" the values of a particular column of a matrix A.

Proposition 2
Let $A = \begin{bmatrix}\vec{a}_1 & \cdots & \vec{a}_n\end{bmatrix}$ be an $m \times n$ matrix and let $\vec{e}_i$ denote the vector in $\mathbb{R}^n$ whose coordinates are all equal to zero, except for the i-th coordinate, which is equal to 1. Then

$$A\vec{e}_i = \vec{a}_i$$

That is, the result of matrix-vector multiplication $A\vec{e}_i$ is the i-th column $\vec{a}_i$ of A.

The following example shows how the "extraction" works.

Example 7

Let $A = \begin{bmatrix}1 & -2 & 3 & 0\\ 0 & 1 & 2 & 7\\ -1 & 2 & 2 & 3\end{bmatrix}$ and $\vec{e}_2 = \begin{bmatrix}0\\ 1\\ 0\\ 0\end{bmatrix} \in \mathbb{R}^4$. Note that the **second** coordinate of $\vec{e}_2$ is 1 , while all other coordinates are 0. Then

$$A\vec{e}_2 = \begin{bmatrix}1 & -2 & 3 & 0\\ 0 & 1 & 2 & 7\\ -1 & 2 & 2 & 3\end{bmatrix}\begin{bmatrix}0\\ 1\\ 0\\ 0\end{bmatrix} = 0\begin{bmatrix}1\\ 0\\ -1\end{bmatrix} + 1\begin{bmatrix}-2\\ 1\\ 2\end{bmatrix} + 0\begin{bmatrix}3\\ 2\\ 2\end{bmatrix} + 0\begin{bmatrix}0\\ 7\\ 3\end{bmatrix} = \begin{bmatrix}-2\\ 1\\ 2\end{bmatrix}$$

which is precisely the **second** column of A.

Exercise 2

A question appears in Mobius

Connection with Linear Systems

Linear Systems and Matrix-Vector Multiplication

It is rather useful to look at systems of linear equations from the perspective of matrix-vector multiplication.

Example 1

Consider a linear system

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= -7 \\ -x_1 - 4x_2 + 3x_3 &= 8 \end{aligned}$$

Let

$$A = \begin{bmatrix} 1 & 3 & -2 \\ -1 & -4 & 3 \end{bmatrix} \quad , \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad , \quad \vec{b} = \begin{bmatrix} -7 \\ 8 \end{bmatrix}$$

Furthermore, let

$$\vec{a}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad , \quad \vec{a}_2 = \begin{bmatrix} 3 \\ -4 \end{bmatrix} \quad , \quad \vec{a}_3 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

be the columns of A so that $A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix}$. Now we can write

$$\begin{bmatrix} x_1 + 3x_2 - 2x_3 \\ -x_1 - 4x_2 + 3x_3 \end{bmatrix} = \begin{bmatrix} -7 \\ 8 \end{bmatrix}$$

which can be rewritten as

$$x_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ -4 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} -7 \\ 8 \end{bmatrix}$$

In turn, this is equivalent to

$$x_1\,\vec{a}_1 + x_2\,\vec{a}_2 + x_3\,\vec{a}_3 = \vec{b}.$$

Notice that the left-hand side of this equation is equal to $A\vec{x}$. Our system above can be expressed as

$$\begin{bmatrix} 1 & 3 & -2 \\ -1 & -4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -7 \\ 8 \end{bmatrix}$$

or more simply as

$$A\,\vec{x} = \vec{b}$$

This shows that we can rewrite a system of linear equations in a much more compact form using matrix-vector multiplication.

Exercise

A question appears in Mobius

Example 2

Observe that

$$\begin{bmatrix} 1 & 5 & -1 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} - 2\begin{bmatrix} 5 \\ -2 \end{bmatrix} + 1\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -7 \\ 13 \end{bmatrix}$$

It follows from the given form that $\vec{x} = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$ is a solution to the system

$$\begin{aligned} x_1 + 5x_2 - x_3 &= -7 \\ 2x_1 - 2x_2 + x_3 &= 13 \end{aligned}$$

Remarks

1. Every system of linear equations can be written as $A\vec{x} = \vec{b}$.
2. The system $A\vec{x} = \vec{b}$ is consistent if and only if $\vec{b}$ can be expressed as a linear combination of the columns of A.
3. If $\vec{a}_1 \dots \vec{a}_n$ are the columns of a $m \times n$ matrix and $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, then $\vec{x}$ satisfies $A\vec{x} = \vec{b}$ if and only if $x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{b}$.

Matrices Equal Theorem

Writing systems of linear equations as a matrix-vector multiplication also reveals some interesting properties about matrices. As the next example illustrates, matrices can sometimes behave quite differently from real numbers.

Example 3

Let

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, \; B = \begin{bmatrix} 3 & -1 \\ 2 & 3 \end{bmatrix}, \; \vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Then

$$A\,\vec{x} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

$$B\,\vec{x} = \begin{bmatrix} 3 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

Thus, $A\,\vec{x} = B\,\vec{x}$ with $\vec{x} \neq \vec{0}$, but $A \neq B$!

This might seem strange because we are used to working with real numbers a, b, $x \in \mathbb{R}$ where, if $ax = bx$ and $x \neq 0$, it means that $a = b$. However, **matrices do not satisfy this property.**

Using this observation, we can derive the following theorem, which tells us under which condition it is possible to conclude that the two matrices A and B are equal. This result is very useful when studying linear transformations in $\mathbb{R}^n$.

Theorem 3: Matrices Equal Theorem
Let A and B be $m \times n$ matrices. If $A\vec{x} = B\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$, then $A = B$.

3.4 - Matrix Multiplication

Matrix Multiplication

Definition of Matrix Multiplication

Let's start with the definition of matrix multiplication.

Definition: Matrix Multiplication as a Linear Combination of Columns
If A is an $m \times n$ matrix and $B = [\vec{b}_1 \dots \vec{b}_k]$ is an $n \times k$ matrix, then the **product** AB is the $m \times k$ matrix $AB = \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_k \end{bmatrix}$.

Example 1

Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 \\ 1 & -1 \\ 2 & 2 \end{bmatrix}$$

so

$$\vec{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \vec{b}_2 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$$

Then

$$A\vec{b}_1 = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$$

$$A\vec{b}_2 = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$$

We conclude that

$$AB = [A\vec{b}_1 \ A\vec{b}_2] = \begin{bmatrix} 9 & 6 \\ 0 & 1 \end{bmatrix}$$

In the above example, notice that $(A_{2\times 3})(B_{3\times 2}) = (AB)_{2\times 2}$. That is, the number of columns of A is equal to the number of rows of B.

Remark
In general,

$$(A_{m\times n})(B_{n\times k}) = (AB)_{m\times k}$$

That is, the matrix multiplication is defined whenever the number of columns of A is equal to the number of rows of B. In this case, the number of rows of AB is equal to the number of rows of A, and the number of columns of AB is equal to the number of columns of B.

Example 2

Is the product

$$AB = \begin{bmatrix} 3 & 6 \\ 9 & 8 \end{bmatrix} \begin{bmatrix} 4 & 6 & 3 \\ 0 & 6 & 1 \end{bmatrix}$$

defined?

Solution

The product is defined because the number of columns of A is equal to the number of rows of B.

Example 3

Is the product

$$AB = \begin{bmatrix} 1 & 1 & 3 \\ 4 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

defined?

Solution

The product is not defined because the number of columns of A is not equal to the number of rows of B.

Exercise 1

A question appears in Mobius

We see that computing matrix products can be tedious if done as in Example 1 above. Normally, we use dot products to compute them more efficiently.

Definition: Matrix Multiplication as a Dot Product of Rows and Columns
Let A be an $m \times n$ matrix with rows $\vec{r}_1, \ldots, \vec{r}_m \in \mathbb{R}^n$. That is

$$A = \begin{bmatrix} \vec{r}_1^T \\ \vdots \\ \vec{r}_m^T \end{bmatrix}$$

Let B be an $n \times k$ matrix with columns $\vec{b}_1, \ldots, \vec{b}_k \in \mathbb{R}^n$. That is

$$B = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_k \end{bmatrix}$$

We define the **product** AB as $(AB)_{ij} = \vec{r}_i \cdot \vec{b}_j$, where $\vec{r}_i \cdot \vec{b}_j$ denotes the dot product of the ith row of A and the jth column of B. In other words

$$AB = \begin{bmatrix} \vec{r}_1^T \\ \vdots \\ \vec{r}_m^T \end{bmatrix} \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_k \end{bmatrix} = \begin{bmatrix} \vec{r}_1 \cdot \vec{b}_1 & \cdots & \vec{r}_1 \cdot \vec{b}_k \\ \vdots & \ddots & \vdots \\ \vec{r}_m \cdot \vec{b}_1 & \cdots & \vec{r}_m \cdot \vec{b}_k \end{bmatrix}$$

Example 4

Compute the product

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 \\ 4 & -2 & 1 \end{bmatrix}$$

Solution

Define

$$\vec{r}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \vec{r}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \vec{b}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \vec{b}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \vec{b}_3 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Then

$$\begin{aligned} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 \\ 4 & -2 & 1 \end{bmatrix} &= \begin{bmatrix} \vec{r}_1 \cdot \vec{b}_1 & \vec{r}_1 \cdot \vec{b}_2 & \vec{r}_1 \cdot \vec{b}_3 \\ \vec{r}_2 \cdot \vec{b}_1 & \vec{r}_2 \cdot \vec{b}_2 & \vec{r}_2 \cdot \vec{b}_3 \end{bmatrix} \\ &= \begin{bmatrix} 1(1)+2(4) & 1(1)+2(-2) & 1(3)+2(1) \\ 3(1)+4(4) & 3(1)+4(-2) & 3(3)+4(1) \end{bmatrix} \\ &= \begin{bmatrix} 9 & -3 & 5 \\ 19 & -5 & 13 \end{bmatrix} \end{aligned}$$

The next example shows how to compute the product of matrices, using both definitions of matrix multiplication.

A slideshow appears in Mobius.

Slide

Example 5

Compute the product $AB = \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 3 \\ -1 & 1 \\ 2 & 0 \end{bmatrix}$ using both definitions of matrix multiplication.

Solution (Matrix multiplication as a linear combination of columns)
Let

$$\vec{b}_1 = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}, \quad \vec{b}_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

Then, according to the first definition of matrix multiplication, $AB = A \begin{bmatrix} \vec{b}_1 & \vec{b}_2 \end{bmatrix} = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 \end{bmatrix}$.

We use matrix-vector multiplication to find $A\vec{b}_1$:

$$\begin{aligned} A\vec{b}_1 &= \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} \\ &= 0 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + (-1) \begin{bmatrix} 0 \\ -1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 1 \end{bmatrix} \end{aligned}$$

We use matrix-vector multiplication to find $A\vec{b}_2$:

$$\begin{aligned} A\vec{b}_2 &= \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \\ &= 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ -1 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 5 \end{bmatrix} \end{aligned}$$

Combining these two results, we find that $AB = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 5 \end{bmatrix}$.

Slide

Example 5 Continued

Compute the product $AB = \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 3 \\ -1 & 1 \\ 2 & 0 \end{bmatrix}$ using both definitions of matrix multiplication.

Solution (Matrix multiplication as a dot product of rows and columns)
Define

$$\vec{r}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{r}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{b}_1 = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}, \quad \vec{b}_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

Then it follows from the dot product definition of matrix multiplication that

$$\begin{aligned} AB &= \begin{bmatrix} \vec{r}_1 \cdot \vec{b}_1 & \vec{r}_1 \cdot \vec{b}_2 \\ \vec{r}_2 \cdot \vec{b}_1 & \vec{r}_2 \cdot \vec{b}_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \cdot 0 + 0 \cdot (-1) + 1 \cdot 2 & 1 \cdot 3 + 0 \cdot 1 + 1 \cdot 0 \\ 2 \cdot 0 + (-1) \cdot (-1) + 0 \cdot 2 & 2 \cdot 3 + (-1) \cdot 1 + 0 \cdot 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 3 \\ 1 & 5 \end{bmatrix} \end{aligned}$$

Exercise 2

A question appears in Mobius

Some Additional Characteristics of Matrix Multiplication

The next example shows an important way in which matrix multiplication differs from multiplication of real numbers.

Example 6

Given

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$$

compute AB and BA.

Solution

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \\ BA &= \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

Notice that $AB \neq BA$, even though AB and BA are both defined and the same size!

This is different from our experience with real numbers, where the order of multiplication does not matter, as ab is always equal to ba for $a, b, \in \mathbb{R}$.

We can also combine matrix multiplication with other matrix operations, such as transposing. This is shown in the next example.

Example 7

Given

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$$

compute $(AB)^T$ and A^TB^T.

Solution

$$(AB)^T = \begin{bmatrix} -1 & 5 \\ -1 & 11 \end{bmatrix}^T = \begin{bmatrix} -1 & -1 \\ 5 & 11 \end{bmatrix}$$
$$A^TB^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 6 & 6 \end{bmatrix}$$

Notice that $(AB)^T \neq A^TB^T$.

However, if we now compute B^TA^T :

$$B^TA^T = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 5 & 11 \end{bmatrix}$$

We see that $(AB)^T = B^TA^T$.

Exercise 3

A question appears in Mobius

Properties of Matrix Multiplication

In the following theorem we summarize the properties of matrix multiplication.

Theorem 1: Properties of Matrix Multiplication
Let $k \in \mathbb{R}$, A, B, and C be matrices such that the following are defined.

1. $IA = A$ (I is an identity matrix)
2. $A(BC) = (AB)C$ (associative law)
3. $A(B + C) = AB + AC$ (left distributive law)
4. $(B + C)A = BA + CA$ (right distributive law)
5. $k(AB) = (kA)B = A(kB)$
6. $(AB)^T = B^T A^T$

Remarks

- Property 2 holds also for the matrix-vector product $A(B\vec{x}) = (AB)\vec{x}$.
- Property 6 generalizes to $(A_1 A_2 \ldots A_k)^T = A_k^T \ldots A_2^T A_1^T$, and as a consequence of this generalization, for any positive integer k, $(A^k)^T = (A^T)^k$.

Example 8

Simplify the expression $A + A(B + I)$ using properties of matrix multiplication.

Solution

We have

$$\begin{aligned}
A + A(B + I) &= A + AB + AI && \text{by property 3} \\
&= A + AB + A && \text{by property 1} \\
&= (A + A) + AB && \text{by properties of matrix addition} \\
&= 2A + AB
\end{aligned}$$

Example 9

Simplify the expression $A(3B - C) + (A - 2B)C + 2B(C + 2A)$ using properties of matrix multiplication.

Solution

We have

$$\begin{aligned}
A(3B - C) + (A - 2B)C + 2B(C + 2A) &= 3AB - AC + AC - 2BC + 2BC + 4BA \\
&= 3AB + 4BA
\end{aligned}$$

Therefore the answer is $3AB + 4BA$.

Note that

- $(A - 2B)C = AC - 2BC$ (C must remain on the right)
- $3AB + 4BA \neq 7AB$ (since we cannot conclude that $AB = BA$)

Exercise 4

A question appears in Mobius

3.5 - Matrix Inverse

Matrix Inverse

Definition of the Inverse of a Matrix

Let's start with the definition of the inverse of a matrix.

Definition
Let A be an $n \times n$ matrix. If there exists an $n \times n$ matrix B such that $AB = I = BA$, then A is **invertible** and B is an **inverse** of A (and B is invertible with A on the inverse of B).

Example 1

Let

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

Is B an inverse of A?

Solution

Note that

$$AB = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = BA$$

We conclude that B is the inverse of A. In particular, this means that A is invertible.

Example 2

Given $A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$, does there exist a matrix $B = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$, where $b_1, b_2, b_3, b_4 \in \mathbb{R}$, such that B is an inverse of A?

Solution

Computing AB gives

$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} b_1 + 2b_3 & b_2 + 2b_4 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

So A is not invertible. Notice that A is non-zero.

Exercise 1

A question appears in Mobius

Recall that the **rank** of a matrix is the number of leading entries in any REF of the matrix.

Theorem 1: Rank of a Square Matrix and of Its Inverse
Let A and B be $n \times n$ matrices such that $AB = I$. Then $BA = I$. Moreover, $\text{rank}(A) = \text{rank}(B) = n$.

It follows from the above theorem that to show B is an inverse of A, we only need to demonstrate that $AB = I$ or $BA = I$. We also deduce that if A is an invertible $n \times n$ matrix, then $\text{rank}(A) = n$. So the RREF of A is I, the $n \times n$ identity matrix.

Exercise 2

Explain why the RREF of an $n \times n$ matrix A with $\text{rank}(A) = n$ is the $n \times n$ identity matrix.

A question appears in Mobius

Theorem 2: Uniqueness of Inverse
Let A be an $n \times n$ matrix. If B and C are $n \times n$ matrices and are both inverses of A, then $B = C$. This means that the inverse of A is unique.

Proof

Let A, B, C be $n \times n$ matrices, such that A is invertible and B, C are inverses of A. Thus $AB = I = BA$ and $AC = I = CA$.

Thus
$$B = BI = B(AC) = (BA)C = IC = C.$$

□

We now have that if A has an inverse, then the inverse is unique, and we denote this unique inverse by A^{-1}.

Theorem 3: Properties of the Inverse
Let A and B be real $n \times n$ matrices and be invertible and let $t \in \mathbb{R}$, $t \neq 0$. Then

1. $(tA)^{-1} = \frac{1}{t}A^{-1}$.
2. $(AB)^{-1} = B^{-1}A^{-1}$, in general $(A_1A_2 \dots A_k)^{-1} = A_k^{-1} \dots A_2^{-1}A_1^{-1}$ for any invertible $n \times n$ matrix A_i.
3. $(A^k)^{-1} = (A^{-1})^k$ for k a positive integer. Here A^k denotes the product $A \cdot A \cdots A$, where the multiplication is repeated k times.
4. $(A^T)^{-1} = (A^{-1})^T$.
5. $(A^{-1})^{-1} = A$

Proof of Property 2

In order to show that $B^{-1}A^{-1}$ is the inverse of AB, we need to show that their product is the identity matrix
$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

So
$$(AB)^{-1} = B^{-1}A^{-1}$$

□

Proof of Property 4

In order to show that $(A^{-1})^T$ is the inverse of A^T, we need to show that their product is the identity matrix

$$(A^T)(A^{-1})^T = (A^{-1}A)^T = I^T = I$$

So

$$(A^T)^{-1} = (A^{-1})^T$$

□

Example 3

Let $A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$ and $t = 3$. Determine $(tA)^{-1}$.

Solution

By property 1 of the theorem above, we know that $(tA)^{-1} = \dfrac{1}{t}A^{-1}$ for $t \neq 0$. In this case, we have $t = 3$ and $A^{-1} = \begin{bmatrix} 2/3 & -1/3 \\ 1/6 & 1/6 \end{bmatrix}$, hence

$$(3A)^{-1} = \frac{1}{3}A^{-1} = \frac{1}{3}\begin{bmatrix} 2/3 & -1/3 \\ 1/6 & 1/6 \end{bmatrix} = \begin{bmatrix} 2/9 & -1/9 \\ 1/18 & 1/18 \end{bmatrix}$$

Calculation of the Inverse of a Matrix

Given an $n \times n$ matrix A, we want to know whether A is invertible and, if so, how to compute A^{-1}. We will study this procedure by considering a 3×3 matrix A. Our method generalizes to $n \times n$ matrices.

Let A be a 3×3 matrix. If A is invertible, then there exists a 3×3 matrix X such that

$$AX = I$$

Let $\vec{x}_1, \vec{x}_2$ and $\vec{x}_3$ be the columns of X so that $X = [\vec{x}_1\ \vec{x}_2\ \vec{x}_3]$. Notice that the columns of the 3×3 identity matrix I are the vectors $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ which we have seen in the lesson on matrix-vector multiplication.

We can write the equation $AX = I$ as

$$A[\vec{x}_1\ \vec{x}_2\ \vec{x}_3] = [\vec{e}_1\ \vec{e}_2\ \vec{e}_3]$$

Applying matrix-vector multiplication on the left, we obtain

$$[A\vec{x}_1\ A\vec{x}_2\ A\vec{x}_3] = [\vec{e}_1\ \vec{e}_2\ \vec{e}_3]$$

Since the two matrices are equal, their respective columns must be equal, therefore

$$A\vec{x}_1 = \vec{e}_1, \quad A\vec{x}_2 = \vec{e}_2, \quad A\vec{x}_3 = \vec{e}_3$$

We have three systems of equations with the same coefficient matrix, A, so we construct an augmented matrix

$$[A \mid \vec{e}_1\ \vec{e}_2\ \vec{e}_3] = [A \mid I]$$

and solve them simultaneously. We consider two cases:

- **Case 1:** the RREF of A is I. The
$$[A \mid I] \sim [I \mid B] = \left[I \mid \vec{b}_1\ \vec{b}_2\ \vec{b}_3\right]$$
where $B = [\vec{b}_1\ \vec{b}_2\ \vec{b}_3]$ is the 3×3 matrix obtained from I via the same row operations that carry A to I. The symbol "$\sim$" means that there is a sequence of elementary row operations that takes the matrix $[A \mid I]$ to the matrix $[I \mid B]$.
Thus our reduced systems are
$$I\vec{x}_1 = \vec{b}_1, \quad I\vec{x}_2 = \vec{b}_2, \quad I\vec{x}_3 = \vec{b}_3$$
so $\vec{x}_1 = \vec{b}_1,\ \vec{x}_2 = \vec{b}_2$, and $\vec{x}_3 = \vec{b}_3$. Hence
$$X = [\vec{x}_1\ \vec{x}_2\ \vec{x}_3] = [\vec{b}_1\ \vec{b}_2\ \vec{b}_3] = B$$
So
$$AX = AB = [A\vec{b}_1\ A\vec{b}_2\ A\vec{b}_3] = [\vec{e}_1\ \vec{e}_2\ \vec{e}_3] = I$$
Since $AB = I$, $B = A^{-1}$.
- **Case 2:** the RREF of A is not I. This means that $\text{rank}(A) < 3$, so at least one of the systems $A\vec{x}_1 = \vec{e}_1$, $A\vec{x}_2 = \vec{e}_2$ and $A\vec{x}_3 = \vec{e}_3$ has no solutions (do you see why?). Consequently, at least one of the three vectors $\vec{x}_1, \vec{x}_2, \vec{x}_3$ does not exist, and so A is not invertible.

Example 4

Is the matrix $A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$ invertible?

Solution

We have
$$\left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right] \underset{R_2-2R_1}{\sim} \left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 0 & -1 & -2 & 1 \end{array}\right] \underset{R_1+3R_2}{\sim} \left[\begin{array}{cc|cc} 2 & 0 & -5 & 3 \\ 0 & -1 & -2 & 1 \end{array}\right] \underset{\substack{(1/2)R_1 \\ -R_2}}{\sim}$$
$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{-5}{2} & \frac{3}{2} \\ 0 & 1 & 2 & -1 \end{array}\right]$$

So A is invertible and $A^{-1} = \begin{bmatrix} \frac{-5}{2} & \frac{3}{2} \\ 2 & -1 \end{bmatrix}$.

Exercise 3

For $A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$ and $A^{-1} = \begin{bmatrix} \frac{-5}{2} & \frac{3}{2} \\ 2 & -1 \end{bmatrix}$, verify that $AA^{-1} = I$.

A question appears in Mobius

Remark
Finding A^{-1} by row reducing $[A|I]$ to $[I|A^{-1}]$ is called the **matrix inversion algorithm**.

Let's see an example of the **matrix inversion algorithm** in action.

Example 5

Is the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ invertible?

Solution

We have

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array}\right] \underset{R_2-2R_1}{\sim} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right]$$

The RREF of A is

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

So A is not invertible (note also that $\text{rank}(A) = 1 < 2$).

Example 6

Given $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & -2 \\ 1 & 2 & -2 \end{bmatrix}$, find A^{-1}, if it exists.

Solution

We have

$$A = \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 1 & 1 & -2 & 0 & 1 & 0 \\ 1 & 2 & -2 & 0 & 0 & 1 \end{array}\right] \underset{\substack{R_2-R_1\\R_3-R_1}}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 2 & -1 & -1 & 0 & 1 \end{array}\right] \underset{R_3-2R_2}{\sim}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right] \underset{\substack{R_1+R_3\\R_2+R_3}}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -2 & 1 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right]$$

So

$$A^{-1} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & -1 & 1 \\ 1 & -2 & 1 \end{bmatrix}$$

Exercise 4

For $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & -2 \\ 1 & 2 & -2 \end{bmatrix}$ and $A^{-1} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & -1 & 1 \\ 1 & -2 & 1 \end{bmatrix}$, verify that $AA^{-1} = I$.

A question appears in Mobius

Exercise 5

A question appears in Mobius

Properties of the Inverse of a Matrix

Now that we know how to calculate the inverse of a matrix, we will explore its properties with respect to various operations on matrices.

Proposition 4: Cancellation Laws
Let A be an invertible $n \times n$ matrix.

1. For all $n \times k$ matrices B, C , if $AB = AC$, then $B = C$ (left cancellation).
2. For all $k \times n$ matrices D, E, if $DA = EA$, then $D = E$ (right cancellation).

Proof of Proposition 1

We have

$$\begin{aligned} AB &= AC \\ A^{-1}(AB) &= A^{-1}(AC) && \text{since } A \text{ is invertible} \\ (A^{-1}A)B &= (A^{-1}A)C \\ IB &= IC \\ B &= C \end{aligned}$$

□

Note that our two cancellation laws require A to be invertible. In the following example, the matrix $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is not invertible.

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

Despite the fact that the above equality holds, the matrices $B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ are distinct.

Example 7

If A, B, C are $n \times n$ matrices such that A is invertible and $AB = CA$, does $B = C$?

Solution

No. Consider

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}$$

Then A is invertible and

$$AB = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = CA$$

So $AB = CA$ but $B \neq C$.

Note that since A is invertible, from $AB = CA$ we obtain $B = A^{-1}CA$. As $B \neq C$, we see that $A^{-1}CA \neq C$.

It might be tempting to try to simplify the expression $A^{-1}CA$ by swapping C and A (or A^{-1} and C), but it is important to remember that, in general, this is not allowed for matrix multiplication.

Exercise 6

For A, B $n \times n$ matrices with A, B and $A + B$ invertible, does $(A + B)^{-1} = A^{-1} + B^{-1}$?

A question appears in Mobius

There are many different perspectives on invertibility. We summarize some of them in the invertible matrix theorem. We will expand this theorem further in later lessons.

Theorem 5: The Invertible Matrix Theorem
Let A be an $n \times n$ matrix. The following statements are equivalent. That is, if one of the following properties holds, then so do all of the others; conversely, if one of the properties fails, then so do all of the others.

1. A is invertible
2. $\text{rank}(A) = n$
3. The RREF of A is $\mathbb{I}$
4. For all $\vec{b} \in \mathbb{R}^n$, the system $A\vec{x} = \vec{b}$ is consistent and has a unique solution
5. A^T is invertible
6. $\text{rank}(A^T) = n$
7. The RREF of A^T is $\mathbb{I}$
8. For all $\vec{b} \in \mathbb{R}^n$, the system $A^T\vec{x} = \vec{b}$ is consistent and has a unique solution

As the invertible matrix theorem shows, knowing that A is invertible tells us a lot of information about A, A^T, and their associated systems of equations. In particular, for a system of equations $A\vec{x} = \vec{b}$ with A invertible, we have

$$\begin{aligned} A\vec{x} &= \vec{b} \\ A^{-1}(A\vec{x}) &= A^{-1}\vec{b} \\ (A^{-1}A)\vec{x} &= A^{-1}\vec{b} \\ I\vec{x} &= A^{-1}\vec{b} \\ \vec{x} &= A^{-1}\vec{b} \end{aligned}$$

This means that we can solve a system of this type directly by calculating $A^{-1}\vec{b}$ and that this is a unique solution as predicted by the invertible matrix theorem.

Example 8

Consider the system of equations $A\vec{x} = \vec{b}$ with

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

Find the solution of this system.

Solution

A is invertible (we computed A^{-1} in an eariler example, by row reducing $[A|I] \sim [I|A^{-1}]$), so

$$\vec{x} = A^{-1}\vec{b} = \begin{bmatrix} \frac{-5}{2} & \frac{3}{2} \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{-23}{2} \\ 9 \end{bmatrix}$$

is the unique solution.

Alternatively, to solve $A\vec{x} = \vec{b}$, we could row reduce:

$$[A|\vec{b}] \sim \left[\begin{array}{cc|c} 1 & 0 & -\frac{23}{2} \\ 0 & 1 & 9 \end{array}\right]$$

In either case, we use the same row operations.

Exercise 7

A question appears in Mobius

Unit 4

Independence and Basis

4.1 - Linear Combinations and Span

Linear Combinations and Span

Spanning Sets

In Unit 1 we learned about the notion of **linear combinations** of vectors $\vec{v}_1, \ldots, \vec{v}_k$. In this lesson we will take a closer look at **spanning sets**, which are sets that contain all possible linear combinations of $\vec{v}_1, \ldots, \vec{v}_k$.

Definition
Let $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ be a set of vectors in $\mathbb{R}^n$. The **span** of B is the set which contains all of the linear combinations of the elements of B.
We write this as

$$\operatorname{Span} B = \{c_1\vec{v}_1 + \cdots + c_k\vec{v}_k \mid c_1, \ldots, c_k \in \mathbb{R}\}$$

We can state the relationship between B and $\operatorname{Span} B$ in any of the following ways:

- The set $\operatorname{Span} B$ is **spanned** by B.
- B is a **spanning set** for $\operatorname{Span} B$.

Example 1

Let $B = \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$. Describe the set $\operatorname{Span} B$.

Solution

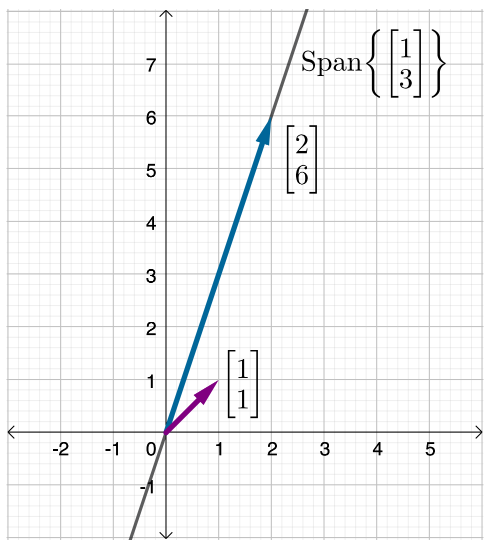
The set $\operatorname{Span} B = \left\{ c \begin{bmatrix} 1 \\ 3 \end{bmatrix} \mid c \in \mathbb{R} \right\}$ consists of all the linear combinations of the vector $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$. Notice that in this case, linear combinations can only be scalar multiples of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

For example, the vector $\begin{bmatrix} 2 \\ 6 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is an element of $\operatorname{Span} B$, while the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is not an element of $\operatorname{Span} B$

because it is not a scalar multiple of $\begin{bmatrix}1\\3\end{bmatrix}$.

Using the familiar set notation, we write $\begin{bmatrix}2\\6\end{bmatrix} \in \operatorname{Span} B$ to indicate that $\begin{bmatrix}2\\6\end{bmatrix}$ is an element of $\operatorname{Span} B$ and $\begin{bmatrix}1\\1\end{bmatrix} \notin \operatorname{Span} B$ to indicate that $\begin{bmatrix}1\\1\end{bmatrix}$ is not an element of $\operatorname{Span} B$.

Geometrically, $\operatorname{Span} B$, or in other words $\operatorname{Span}\left\{\begin{bmatrix}1\\3\end{bmatrix}\right\}$, is represented by a line in $\mathbb{R}^2$. Notice that the vector $\begin{bmatrix}2\\6\end{bmatrix}$ lies along the line, but the vector $\begin{bmatrix}1\\1\end{bmatrix}$ does not.



Example 2

Check if $\begin{bmatrix}2\\3\end{bmatrix} \in \operatorname{Span}\left\{\begin{bmatrix}4\\5\end{bmatrix}, \begin{bmatrix}3\\3\end{bmatrix}\right\}$.

Solution

Let $c_1, c_2 \in \mathbb{R}$ and consider

$$\begin{bmatrix}2\\3\end{bmatrix} = c_1\begin{bmatrix}4\\5\end{bmatrix} + c_2\begin{bmatrix}3\\3\end{bmatrix} = \begin{bmatrix}4c_1+3c_2\\5c_1+3c_2\end{bmatrix}.$$

We obtain the system of equations

$$\begin{aligned}2 &= 4c_1 + 3c_2\\ 3 &= 5c_1 + 3c_2\end{aligned}$$

Subtracting the first equation from the second one gives $c_1 = 1$ and it follows from the first equation that $2 = 4(1) + 3c_2$, so $c_2 = -2/3$. We see that

$$\begin{bmatrix}2\\3\end{bmatrix} = 1\begin{bmatrix}4\\5\end{bmatrix} - \frac{2}{3}\begin{bmatrix}3\\3\end{bmatrix}$$

so

$$\begin{bmatrix}2\\3\end{bmatrix} \in \operatorname{Span}\left\{\begin{bmatrix}4\\5\end{bmatrix}, \begin{bmatrix}3\\3\end{bmatrix}\right\}$$

Example 3

Check if $\begin{bmatrix}1\\2\\3\end{bmatrix} \in \text{Span}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$.

Solution

Let $c_1, c_2 \in \mathbb{R}$ and consider

$$\begin{bmatrix}1\\2\\3\end{bmatrix} = c_1\begin{bmatrix}1\\0\\1\end{bmatrix} + c_2\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}c_1+c_2\\c_2\\c_1\end{bmatrix}$$

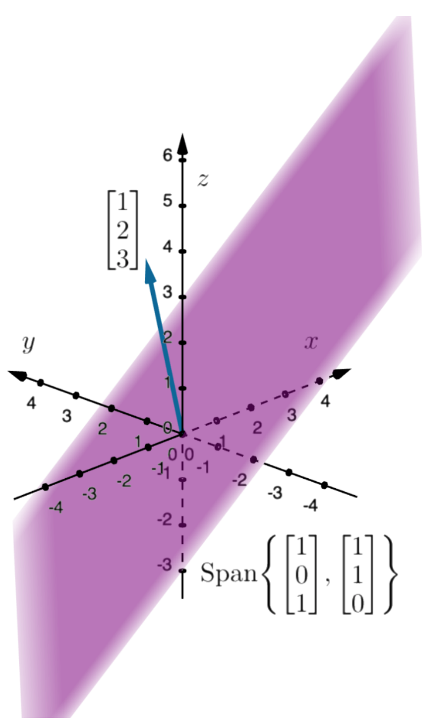
We obtain

$$\begin{aligned}1 &= c_1 + c_2\\2 &= c_2\\3 &= c_1\end{aligned}$$

Clearly, $c_1 = 3$ and $c_2 = 2$, but $c_1 + c_2 = 3 + 2 = 5 \neq 1$, so our system has no solution. Thus

$$\begin{bmatrix}1\\2\\3\end{bmatrix} \notin \text{Span}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$$

Geometrically, $\text{Span}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$ is represented by a plane in $\mathbb{R}^3$. Notice that the vector $\begin{bmatrix}1\\2\\3\end{bmatrix}$ does not lie on the plane.



For brevity, let $S = \text{Span}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$. Looking back at the previous example, let us see the geometric interpretation of S.

Any $\vec{x} \in S$ is of the form

$$\vec{x} = s\begin{bmatrix}1\\0\\1\end{bmatrix} + t\begin{bmatrix}1\\1\\0\end{bmatrix}, \quad s, t \in \mathbb{R},$$

which is the vector equation of a plane through the origin (since the two vectors in the spanning set are not scalar multiples of one another). Thus any $\vec{x} \in S$ lies on this plane. Also, by definition

$$S = \left\{ s\begin{bmatrix}1\\0\\1\end{bmatrix} + t\begin{bmatrix}1\\1\\0\end{bmatrix} \mid s, t \in \mathbb{R} \right\}$$

represents a plane through the origin.

Exercise 1

A question appears in Mobius

Remark
Recall that a set A is a subset of a set B if every element of A is also an element of B. We write this as $A \subseteq B$. We say that the two sets A and B are equal if they are subsets of each other, and write this as $A = B$.

Example 4

Describe the subset

$$S = \text{Span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$$

of $\mathbb{R}^3$ geometrically.

Solution

By definition,

$$S = \left\{ c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \mid c_1, c_2, c_3 \in \mathbb{R} \right\}$$

so a vector equation for S is

$$\vec{x} = c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad c_1, c_2, c_3 \in \mathbb{R}$$

But

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

so

$$\begin{aligned} \vec{x} &= c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right) \\ &= (c_1 + c_3) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (c_2 + c_3) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\ &= d_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + d_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

where $d_1 = c_1 + c_3$ and $d_2 = c_2 + c_3$. We conclude that S is contained in the x_1x_2-plane of $\mathbb{R}^3$, which we denote by T:

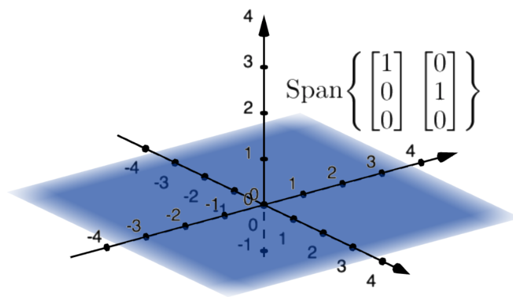
$$S = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\} \subseteq \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} = T$$

We claim that $S = T$. To see that this is the case, we need to prove that $T \subseteq S$. Let $\vec{x} \in T$. Then there exist $d_1, d_2 \in \mathbb{R}$ such that

$$\begin{aligned} \vec{x} &= d_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + d_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\ &= d_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + d_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

so $\vec{x} \in S$ and $T \subseteq S$. Hence $S = T$.

Geometrically, S is the plane $z = 0$ in $\mathbb{R}^3$.



In the example above, we managed to remove one vector from the span without changing it. The following provides a criterion for when this can be done.

Theorem 1
Let $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$. Some vector $\vec{v}_i$ can be written as a linear combination of $\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_k$ if and only if

$$\text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\} = \text{Span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_k\}.$$

Example 5

Suppose that $\text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \text{Span}\{\vec{v}_1, \vec{v}_3\}$. What can you conclude about $\vec{v}_2$?

Solution

If $\text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \text{Span}\{\vec{v}_1, \vec{v}_3\}$, then $\vec{v}_2$ can be written as a linear combination of $\vec{v}_1$ and $\vec{v}_3$. That is, there exist scalars $s, t \in \mathbb{R}$ such that

$$\vec{v}_2 = s\vec{v}_1 + t\vec{v}_3$$

Example 6

Let $A = \left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-2\\-4\end{bmatrix}, \begin{bmatrix}-2\\3\end{bmatrix}\right\}$. Find a subset B of A with the smallest number of elements possible such that $\text{Span}(A) = \text{Span}(B)$.

Solution

Notice that $\begin{bmatrix}-2\\-4\end{bmatrix}$ can be written as a linear combination of $\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-2\\3\end{bmatrix}$:

$$\begin{bmatrix}-2\\-4\end{bmatrix} = -2\begin{bmatrix}1\\2\end{bmatrix} + 0\begin{bmatrix}-2\\3\end{bmatrix}$$

By the previous theorem, we have that

$$\text{Span}\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-2\\-4\end{bmatrix}, \begin{bmatrix}-2\\3\end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-2\\3\end{bmatrix}\right\}$$

Note that we cannot remove any more vectors from the set $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \end{bmatrix} \right\}$ without changing the span.

Therefore $B = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \end{bmatrix} \right\}$ is a subset of A with the smallest number of elements such that $\text{Span}(A) = \text{Span}(B)$.

Remark

In the previous example, $B = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \end{bmatrix} \right\}$ is not the only set we could have chosen. Use the interaction below to convince yourself of this.

Making Connections

Select a checkbox to display the span of the associated vector(s).

External resource: https://www.geogebra.org/material/iframe/id/aazcv7w3/

Example 7

Let $A = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$. Find a subset B of A with the smallest number of elements possible such that $\text{Span}\, A = \text{Span}\, B$ and describe $\text{Span}\, A$ geometrically.

Solution

Choosing vector $v_2 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$ from the set A, we find that

$$\begin{bmatrix} 5 \\ 0 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 4 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

In other words, $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$ can be written as a linear combination of the other vectors in A.

By the theorem, we can remove $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$ from A:

$$\text{Span}\, A = \text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} = \text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Since

$$\begin{bmatrix} 2 \\ 4 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

we can apply the theorem again and remove $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ from A:

$$\text{Span}\, A = \text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} = \text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

As neither $\begin{bmatrix}1\\0\end{bmatrix}$ nor $\begin{bmatrix}0\\1\end{bmatrix}$ are scalar multiples of one another, we cannot reduce further. Thus

$$\operatorname{Span} A = \operatorname{Span}\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$$

and $B = \left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$.

To describe $\operatorname{Span} A$ geometrically, notice that a vector equation for $\operatorname{Span} A$ is

$$\vec{x} = c_1\begin{bmatrix}1\\0\end{bmatrix} + c_2\begin{bmatrix}0\\1\end{bmatrix}, \quad c_1, c_2 \in \mathbb{R}$$

Combining the vectors on the right gives

$$\vec{x} = \begin{bmatrix}c_1\\c_2\end{bmatrix}$$

which means that $\operatorname{Span} A = \mathbb{R}^2$.

Making Connections

Select the tip of the vector $\vec{x}$ and drag it to different locations.

Notice that, no matter which vector $\vec{x}$ you create, it is **always** possible to express $\vec{x}$ as a linear combination of $\begin{bmatrix}1\\3\end{bmatrix}$ and $\begin{bmatrix}-3\\-1\end{bmatrix}$. In other words, $\operatorname{Span}\left\{\begin{bmatrix}1\\3\end{bmatrix},\begin{bmatrix}-3\\-1\end{bmatrix}\right\} = \mathbb{R}^2$.

External resource: https://www.geogebra.org/material/iframe/id/x88txpbr/

Exercise 2

A question appears in Mobius

Exercise 3

A question appears in Mobius

Example 8

Let $S = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$. Prove that $S = \mathbb{R}^3$.

Solution

To prove that $S = \mathbb{R}^3$, we need to show that $\mathbb{R}^3 \subseteq S$ and that $S \subseteq \mathbb{R}^3$.

First, we show that $\mathbb{R}^3 \subseteq S$. Consider

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$$

We first need to show that that $\vec{x} \in S$. To do this, we will show that there exist $c_1, c_2, c_3 \in \mathbb{R}$ such that

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 + c_3 \\ c_2 + 2c_3 \\ c_3 \end{bmatrix}$$

We see that

$$\begin{aligned} x_1 &= c_1 + c_2 + c_3 \\ x_2 &= c_2 + 2c_3 \\ x_3 &= c_3 \end{aligned}$$

From the third equation we find that $c_3 = x_3$. From the second equation we find that

$$c_2 = x_2 - 2c_3 = x_2 - 2x_3$$

From the first equation we find that

$$c_1 = x_1 - c_2 - c_3 = x_1 - (x_2 - 2x_3) - x_3 = x_1 - x_2 + x_3$$

Thus

$$\begin{aligned} c_1 &= x_1 - x_2 + x_3 \\ c_2 &= x_2 - 2x_3 \\ c_3 &= x_3 \end{aligned}$$

so $\vec{x} \in S$.

Since $\vec{x}$ is an arbitrary element of $\mathbb{R}^3$, we conclude that $\mathbb{R}^3 \subseteq S$.

Next, we will show that $S \subseteq \mathbb{R}^3$.

Since

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \in \mathbb{R}^3$$

and $\mathbb{R}^3$ is closed under scalar multiplication and addition (and therefore $\mathbb{R}^3$ is closed under linear combinations),

$$S = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\} \subseteq \mathbb{R}^3$$

Remark

"Closed under scalar multiplication and addition" means that performing these operations on vectors in $\mathbb{R}^3$ will always result in vectors in $\mathbb{R}^3$. In fact, this is true for any vector space $\mathbb{R}^n$.

A question appears in Mobius

A question appears in Mobius

4.2 - Linear Dependence and Linear Independence

Linear Dependence and Linear Independence

Definition of Linear Dependence and Linear Independence

Definition
Let $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ be a set of vectors in $\mathbb{R}^n$. We say that B is **linearly dependent** if there exist $c_1, \ldots, c_k \in \mathbb{R}$, not all zero, such that

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

We say that B is **linearly independent** if the only solution to

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

is $c_1 = \cdots = c_k = 0$, which we call the **trivial solution**.

Example 1

Determine whether the set $\left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\}$ is linearly dependent or linearly independent.

Solution

For $c_1, c_2 \in \mathbb{R}$, we need to verify the solution to the system

$$c_1 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This gives the following system of equations:

$$\begin{aligned} 2c_1 - c_2 &= 0 \\ 3c_1 + 2c_2 &= 0 \end{aligned}$$

Adding twice the first equation to the second one gives $7c_1 = 0$, so $c_1 = 0$ and other equation gives $c_2 = 0$.

Thus, the the only solution to the system is $c_1 = c_2 = 0$, which means that the set $\left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\}$ is linearly independent.

Making Connections

In the following interactive exercise, click and drag the two vectors around to explore when the vectors are linearly dependent and linearly independent.

External resource: https://www.geogebra.org/material/iframe/id/uebb59xn/

From the interactive activity, you may have noticed that two vectors $\vec{x}$ and $\vec{y}$ are linearly dependent only when $\vec{x}$ is a scalar multiple of $\vec{y}$ or $\vec{y}$ is a scalar multiple of $\vec{x}$. However, **this applies only to sets containing two vectors**.

If the set contains 3 or more vectors, we cannot always use scalar multiples as the only way to determine linear dependence.

We can consider why this is the case for 3 (or more) vectors by looking at the set $\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$. Notice that none of these vectors is a scalar multiple of the others, but the set is linearly dependent since

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Example 2

Determine whether the set $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is linearly dependent or linearly independent.

Solution

For $c_1, c_2, c_3 \in \mathbb{R}$ we consider

$$c_1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Then

$$\begin{aligned} c_1 + 2c_2 + c_3 &= 0 \\ c_2 + c_3 &= 0 \\ -c_1 + c_3 &= 0 \end{aligned}$$

From the third equation we find that $c_1 = c_3$.

From the second equation we find that $c_2 = -c_3$.

Hence the first equation becomes

$$\begin{aligned} c_3 + 2(-c_3) + c_3 &= 0 \\ 0 &= 0 \end{aligned}$$

which is true for any $c_3 \in \mathbb{R}$. So, let $c_3 = t \in \mathbb{R}$. Then

$$c_1 = t, \quad c_2 = -t, \quad c_3 = t$$

For any $t \neq 0$, we have non-trivial solutions, so our set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

is linearly dependent.

Remark
When determining whether a set $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent or linearly dependent, we must consider solutions to $c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$ for $c_1, \ldots, c_k \in \mathbb{R}$.

Example 3

Prove that the set $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is linearly independent.

Solution

For $c_1, c_2 \in \mathbb{R}$, consider

$$c_1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

We have

$$\begin{aligned} c_1 + c_2 &= 0 \\ c_2 &= 0 \\ -c_1 + c_2 &= 0 \end{aligned}$$

We see $c_2 = 0$ and it follows that $c_1 = 0$. Hence we have only the trivial solution, so the set is linearly independent.

A question appears in Mobius

A question appears in Mobius

Linear Independence and Span

Finding a Linearly Independent Subset of a Spanning Set

The following theorem gives us a characterization of linear dependence and linear independence.

Theorem 1
A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_k\}$ in $\mathbb{R}^n$ is linearly dependent if and only if

$$\vec{v}_i \in \text{Span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_k\}$$

for some i such that $1 \leq i \leq k$.

This theorem tells us that a linearly independent set is one in which no vector is a linear combination of any other vector in the set. Conversely, if there is a vector in a set that can be written as a linear combination of the others, then the set is linearly dependent.

Algorithm

Given a spanning set, we often want to make it linearly independent without changing the span. Here is how this procedure works:

- If the spanning set is linearly independent, then we cannot remove any vectors without changing the span.
- If the spanning set is linearly dependent, then one of the vectors in the spanning set is a linear combination of the others. We can remove that vector and not change the span. We repeat the process until we have no more dependencies.

Example 1

Consider the set $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ from one of our previous examples. Find a linearly independent subset of S that has the same Span as S.

Solution

In our previous work, we found that, for any real number t,

$$t \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - t \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Setting $t = 1$ and rearranging, we find that

$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

We have written one of the vectors in S as a linear combination of the other two. Thus, by the above theorem, the set S is linearly dependent.

We can therefore remove the vector $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ from S without affecting the span:

$$\text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} = \text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

We cannot reduce the spanning set further since the remaining vectors are linearly independent.

Therefore, a linearly independent subset of S that has the same Span as S is $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Note that, in this example, removing the vector $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ was an arbitrary choice. Since any of the three vectors in S can be written as a linear combination of the other two, we could have removed any one of the vectors in S and still obtained the same span:

$$\begin{bmatrix}1\\0\\-1\end{bmatrix} = \begin{bmatrix}2\\1\\0\end{bmatrix} - \begin{bmatrix}1\\1\\1\end{bmatrix} \quad \Rightarrow \quad \text{Span}\left\{\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}2\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix}2\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$$

$$\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}2\\1\\0\end{bmatrix} - \begin{bmatrix}1\\0\\-1\end{bmatrix} \quad \Rightarrow \quad \text{Span}\left\{\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}2\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}2\\1\\0\end{bmatrix}\right\}$$

Example 2

Prove that the set $\{\vec{v}_1, \dots, \vec{v}_k, \vec{0}\}$ in $\mathbb{R}^n$ is linearly dependent.

Solution

We start by setting up the system $c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{0} = \vec{0}$and seeing if there are non-trivial solutions.

Setting $c_1, \dots, c_k = 0$ and $c_{k+1} = 1$, we have

$$0 \cdot \vec{v}_1 + \cdots + 0 \cdot \vec{v}_k + 1 \cdot \vec{0} = \vec{0}$$

Since we found a non-trivial solution (not all of $c_1, \dots, c_{k+1}$ are zero), we have shown that $\{\vec{v}_1, \dots, \vec{v}_k, \vec{0}\}$ is linearly dependent. Note that any subset of $\mathbb{R}^n$ containing the zero vector will be linearly dependent.

Example 3

Let $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{R}^n$ be such that $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is linearly independent. Prove that $\{\vec{v}_1, \vec{v}_1 + \vec{v}_2, \vec{v}_1 + \vec{v}_2 + \vec{v}_3\}$ is linearly independent.

Solution

Consider

$$c_1\vec{v}_1 + c_2(\vec{v}_1 + \vec{v}_2) + c_3(\vec{v}_1 + \vec{v}_2 + \vec{v}_3) = \vec{0}, \quad c_1, c_2, c_3 \in \mathbb{R}$$

Then we have

$$(c_1 + c_2 + c_3)\vec{v}_1 + (c_2 + c_3)\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$$

Since $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is linearly independent, the only solution to $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3$ is $c_1 = c_2 = c_3 = 0$ and therefore, equating components, we have

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ c_2 + c_3 &= 0 \\ c_3 &= 0 \end{aligned}$$

After solving this system of equations, we find that $c_1 = c_2 = c_3 = 0$, so $\{\vec{v}_1, \vec{v}_1 + \vec{v}_2, \vec{v}_1 + \vec{v}_2 + \vec{v}_3\}$ is linearly independent.

Example 4

For $k \geq 2$, let $\{\vec{v}_1, \dots, \vec{v}_k\}$ be linearly independent. Prove that $\{\vec{v}_1, \dots, \vec{v}_{k-1}\}$ is also linearly independent.

Solution

Assume for a contradiction that $\{\vec{v}_1, \dots, \vec{v}_{k-1}\}$ is linearly dependent. Then there exist $c_1, \dots, c_{k-1} \in \mathbb{R}$, not all zero, such that

$$c_1\vec{v}_1 + \cdots + c_{k-1}\vec{v}_{k-1} = \vec{0}$$

But then, for $\{\vec{v}_1, \dots, \vec{v}_k\}$, we can write:

$$c_1\vec{v}_1 + \cdots + c_{k-1}\vec{v}_{k-1} + 0\vec{v}_k = \vec{0}$$

which shows that $\{\vec{v}_1, \dots, \vec{v}_k\}$ is linearly dependent (remember that not all of the c_i's are zero). This contradicts the initially given fact that $\{\vec{v}_1, \dots, \vec{v}_k\}$ is linearly independent. Hence $\{\vec{v}_1, \dots, \vec{v}_{k-1}\}$ must be linearly independent.

Remark
Every subset of a linearly independent set, including the empty set, is linearly independent.

A question appears in Mobius

A question appears in Mobius

4.3 - Basis

Basis

Definition of Basis

Definition
Let S be a subset of $\mathbb{R}^n$. If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is a linearly independent set of vectors in S such that $S = \text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\}$, then we say that B is a **basis** for S. We define a basis for $\{\vec{0}\}$ to be the empty set, $\varnothing$.

Note that there are two components to the definition of a basis. In order for a set $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$to be a basis for S, we require that

1. B is a **linearly independent set** of vectors in S
2. B is a **spanning set** for S

If B is a basis for S, then **any** vector in S can be built from a linear combination of vectors in B.

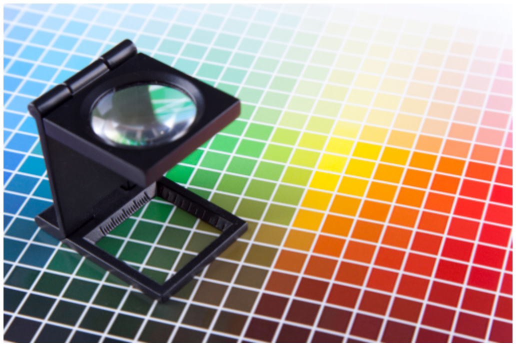
Given a set S, we would like to find a basis B for S (if possible). We can undertand this concept more thoroughly with an analogy, given in the following slideshow.

A slideshow appears in Mobius.

Slide

Intuition Behind Basis: Printing Inks

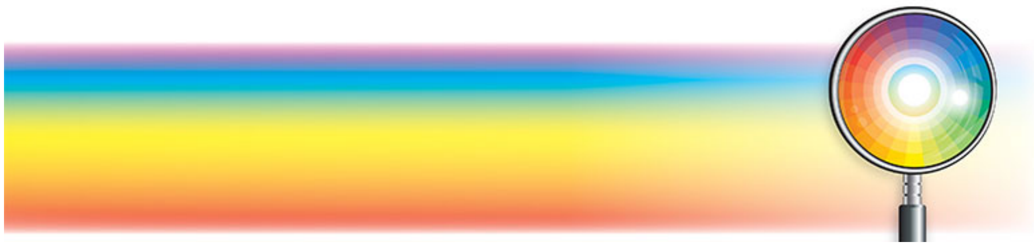
Patrick Daxenbichler/iStock/Getty Images



focusstock/E+/Getty Images

Slide

Intuition Behind Basis: Generating Colours

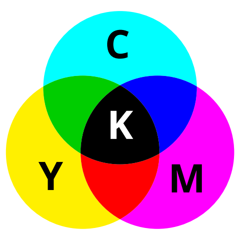


Zoonar RF/Zoonar/Getty Images

We want a set of inks that is:

1. as small as possible (no redundancies), and
2. able to generate all of the colours in the spectrum.

Coprid/iStock/Getty Images



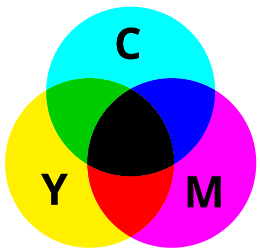
MicrovOne/iStock/Getty Images

Slide

Intuition Behind Basis: Back to the Math

Our set of inks must:

1. be **as small as possible** (no redundancies), and
2. be able to **generate all of the colours** in the spectrum.



MicrovOne/iStock/Getty Images

$B = \{C, M, Y\}$ is a "basis" for the colours in the visible spectrum

A basis B for S must:

1. be a **linearly independent set** of vectors in S
   (no vector is a linear combination of any of the others), and
2. be a **spanning set** for S.

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

B is a basis for $\mathbb{R}^3$

Example 1

Prove that $B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$.

Solution

To show that B is a basis for $\mathbb{R}^2$, we need to show that:

1. B spans $\mathbb{R}^2$, and
2. B is linearly independent.

We will first prove that B spans $\mathbb{R}^2$ or, in other words, $\mathbb{R}^2 = \operatorname{Span} B$. We can do this by showing that any arbitrary element of $\mathbb{R}^2$ can be written as a linear combination of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

For $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$, we have

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

so $\mathbb{R}^2 \subseteq \operatorname{Span} B$. Since $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{R}^2$ and $\mathbb{R}^2$ is closed under linear combinations, $\operatorname{Span} B \subseteq \mathbb{R}^2$ and so $\operatorname{Span} B = \mathbb{R}^2$.

Next, to show that B is linearly independent, let $c_1, c_2 \in \mathbb{R}$ and consider

$$c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This gives us the system

$$\begin{aligned} 1 \cdot c_1 + 0 \cdot c_2 &= 0 \\ 0 \cdot c_1 + 1 \cdot c_2 &= 0 \end{aligned}$$

which has solution

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Since $c_1 = c_2 = 0$, B is linearly independent.

Since B spans $\mathbb{R}^2$ and B is linearly independent, B is a basis for $\mathbb{R}^2$.

Making Connections

This interactive example will help you to see that the set $\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -3 \\ -1 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$.

Drag the vector $\vec{x}$ to various positions, to observe that the set is linearly independent, and that every vector $\vec{x}$ in $\mathbb{R}^2$ can be written as a **unique** linear combination of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ -1 \end{bmatrix}$.

External resource: https://www.geogebra.org/material/iframe/id/pxzjwdtj/

Definition
Let $\vec{e}_i \in \mathbb{R}^n$ be the vector whose i-th entry is 1, and whose other entries are 0. We refer to $\vec{e}_i$ as the ***i*-th standard basis vector** and to the set $\{\vec{e}_1, \ldots, \vec{e}_n\}$ as the **standard basis** of $\mathbb{R}^n$.

Example 2

Determine the standard basis of $\mathbb{R}^2$ and $\mathbb{R}^3$.

Solution

In $\mathbb{R}^2$, $\{\vec{e}_1, \vec{e}_2\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$.

In $\mathbb{R}^3$, $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Example 3

Determine whether $B = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^3$.

Solution

To determine whether B is a basis for $\mathbb{R}^3$, we need to check whether:

1. B spans $\mathbb{R}^3$ and
2. B is linearly independent.

We start by checking whether B spans $\mathbb{R}^3$.

Let $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$ and consider

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad c_1, c_2 \in \mathbb{R}$$

We obtain

$$\begin{array}{rcrcl} c_1 & - & c_2 & = & x_1 \\ 2c_1 & + & 2c_2 & = & x_2 \\ & & c_2 & = & x_3 \end{array}$$

From the third equation we find that $c_2 = x_3$.

From the second equation we find that $c_1 = \dfrac{x_2 - 2x_3}{2}$.

Finally, the first equation gives $\dfrac{x_2 - 2x_3}{2} - x_3 = x_1$, so $2x_1 - x_2 + 4x_3 = 0$.

We see that $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \notin \operatorname{Span} B$ if $2x_1 - x_2 + 4x_3 \neq 0$, so $\mathbb{R}^2 \neq \operatorname{Span} B$ and B is not a basis for $\mathbb{R}^2$.

For example, since $2(1) - 1 + 4(1) \neq 0$, we see that $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \notin \operatorname{Span} B$. This means that B cannot generate all of the vectors in $\mathbb{R}^3$ and therefore cannot span $\mathbb{R}^3$.

Since B does not span $\mathbb{R}^3$, it is not a basis of $\mathbb{R}^3$.

Example 4

Let

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 6 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 5 \\ 7 \\ 12 \\ 17 \end{bmatrix}$$

Let $B = \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ and let $S = \operatorname{Span} B$. Determine whether B is a basis for S.

Solution

To determine whether B is a basis for S, we need to check whether:

1. B spans S, and
2. B is linearly independent.

We already know that $S = \operatorname{Span} B$, since that is how S was defined.

It remains to check whether B is linearly independent or not. Consider

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 + c_4\vec{v}_4 = \vec{0}$$

The resulting system has augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 2 & 1 & 5 & 0 \\ 1 & 2 & 2 & 7 & 0 \\ 2 & 4 & 3 & 12 & 0 \\ 3 & 6 & 4 & 17 & 0 \end{array}\right] \underset{\substack{R_2-R_1 \\ R_3-2R_1 \\ R_4-3R_1}}{\sim} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 5 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 \end{array}\right] \underset{\substack{R_3-R_2 \\ R_4-R_2}}{\sim} \underset{\text{REF}}{\left[\begin{array}{cccc|c} 1 & 2 & 1 & 5 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]}$$

Since the second and the fourth columns have no leading entries, c_2 and c_4 are free variables, so we have non-trivial solutions.

We conclude that B is linearly dependent, so it is not a basis of S.

Exercise 1

A question appears in Mobius

Obtaining a Basis from a Finite Spanning Set

In the previous example, we saw that the set of vectors $B = \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ where

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 6 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 5 \\ 7 \\ 12 \\ 17 \end{bmatrix}$$

was not linearly independent. However, we will see that it is possible to replace B with a subset B' of B, which is linearly independent. Looking at the REF that we had obtained

$$\underset{\text{REF}}{\left[\begin{array}{cccc|c} 1 & 2 & 1 & 5 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]}$$

we see that c_1 and c_3 are the leading variables. Using our previous work, we have that

$$c_1\vec{v}_1 + c_3\vec{v}_3 = \vec{0}$$

which gives the system with augmented matrix

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 0 \\ 3 & 4 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Since there are no free variables, we have only the trivial solution, so $B' = \{\vec{v}_1, \vec{v}_3\}$ is linearly independent.

We claim that $\operatorname{Span} B = \operatorname{Span} B'$. To see that this is the case, we have to show that $\vec{v}_2, \vec{v}_4 \in \operatorname{Span} B'$. We have

$$\vec{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 6 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \end{bmatrix} = 2\vec{v}_1 + 0\vec{v}_3 \in \operatorname{Span} B'$$

To show that $\vec{v}_4 \in \operatorname{Span} B'$, we use row reduction:

$$\left[\begin{array}{cc|c} 1 & 1 & 5 \\ 1 & 2 & 7 \\ 2 & 3 & 12 \\ 3 & 4 & 17 \end{array}\right] \sim \underset{\text{RREF}}{\left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]}.$$

From RREF we deduce that $\vec{v}_4 = 3\vec{v}_1 + 2\vec{v}_3$.

We conclude $S = \text{SpanB} = \operatorname{Span} B'$, and since B' is linearly independent, $B' = \left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \right\}$ is a basis for S.

Example 5

Let

$$B = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -7 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ -9 \end{bmatrix} \right\}$$

Find a basis B' for $\operatorname{Span} B$.

Solution

By construction, the set B spans $\operatorname{Span} B$.

To obtain a basis, we need to find a set of linearly independent vectors in B: the set B' will contain these linearly independent vectors.

Using row reduction, we find that

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ -1 & 2 & 5 & 6 \\ 1 & -3 & -7 & -9 \end{bmatrix} \underset{\substack{R_2+R_1 \\ R_3-R_1}}{\sim} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 3 & 6 & 9 \\ 0 & -4 & -8 & -12 \end{bmatrix} \underset{\substack{1/3R_2 \\ -1/4R_3}}{\sim} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{bmatrix} \underset{\substack{R_1-R_2 \\ R_3-R_2}}{\sim} \underset{\text{RREF}}{\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}}$$

The first two columns have leading entries, so the first two vectors of B comprise B'. Thus the basis B' for $\operatorname{Span} B$ is

$$B' = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} \right\}$$

Further, from RREF we see that

$$\begin{bmatrix} 1 \\ 5 \\ -7 \end{bmatrix} = -1 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$$

and

$$\begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$$

Exercise 2

A question appears in Mobius

4.4 - Coordinates

Coordinates in a Standard Basis

Introduction

Suppose we have a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_k\}$ for a subset S of $\mathbb{R}^n$. The equality $S = \operatorname{Span} \mathcal{B}$ enables us to find a representation for an arbitrary vector $\vec{v} \in S$ as a linear combination of elements from $\mathcal{B}$, while the linear independence guarantees that such a representation

$$\vec{v} = t_1\vec{v}_1 + \cdots + t_k\vec{v}_k$$

is unique. We record this observation in the following theorem.

Theorem 1: Unique Representation Theorem
If $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is a basis for a subset S of $\mathbb{R}^n$, then every $\vec{x} \in S$ can be written as a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$ in a unique way.

From now on, let us assume that $\mathcal{B}$ is a basis for $\mathbb{R}^n$. It turns out that if we know $\mathcal{B}$, we do not need to write the vectors $\vec{v}_1, \ldots, \vec{v}_n$ every time in order to describe $\vec{v} = t_1\vec{v}_1 + \cdots + t_n\vec{v}_n$. The reason is that all the important information is contained in the list of scalars $t_1, \ldots, t_n$.

Definition
Suppose that $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$. If $\vec{v} \in \mathbb{R}^n$ with $\vec{v} = t_1\vec{v}_1 + t_2\vec{v}_2 + \cdots + t_n\vec{v}_n$, then the **coordinate vector** of $\vec{v}$ with respect to the basis $\mathcal{B}$ is

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}$$

We also refer to $[\vec{v}]_{\mathcal{B}}$ as the **coordinates of $\vec{v}$ with respect to $\mathcal{B}$**, or the **$\mathcal{B}$-coordinates of $\vec{v}$.**

While this may seem a bit strange at first glance, this is something that we have secretly been doing all along whenever we work with vectors in $\mathbb{R}^n$.

Example 1

Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ be the standard basis for $\mathbb{R}^3$, and let $\vec{v} = \begin{bmatrix} 2 \\ -5 \\ 7 \end{bmatrix}$.

Find $[\vec{v}]_{\mathcal{B}}$, the coordinate vector of $\vec{v}$ with respect to the basis $\mathcal{B}$.

Solution

We have that $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -5 \\ 7 \end{bmatrix}$ since $\begin{bmatrix} 2 \\ -5 \\ 7 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 5\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 7\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

In general, if $\vec{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, then we have that $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ since $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Vectors, as we know them in $\mathbb{R}^n$, are actually coordinate vectors **with respect to the standard basis**. In other words, even though we use the notation $\vec{v} = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}$, we can also think of this as $\vec{v} = t_1\vec{e}_1 + \cdots + t_n\vec{e}_n$.

Example 2

Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ be the standard basis for $\mathbb{R}^4$, and let $\vec{x} = \begin{bmatrix} 3 \\ -9 \\ -8 \\ 2 \end{bmatrix}$.

Find $[\vec{x}]_{\mathcal{B}}$.

Solution

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ -9 \\ -8 \\ 2 \end{bmatrix}$$

since

$$\begin{bmatrix} 3 \\ -9 \\ -8 \\ 2 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - 9 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} - 8 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

As we can see from the previous examples, working with the standard basis is quite nice since the coordinate vectors can be determined by inspection.

Exercise

A question appears in Mobius

Coordinates in a Non-Standard Basis

Introduction

Instead of always choosing the standard basis for a vector space, we can choose to work with any basis we want. As you can imagine, finding the coordinate vector when using a non-standard basis is not as simple. Our next goal will be to figure out how to do this in a relatively straightforward way.

Remark

Notice that the order of the basis vectors matters. For example, the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is not the same basis for $\mathbb{R}^3$ as the basis $\mathcal{C} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$. As the example below shows, this change in the order of vectors results in different coordinates for our vectors.

Example 1

Let $\mathcal{E} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ be the standard basis for $\mathbb{R}^3$, let $\mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ be another basis for $\mathbb{R}^3$, and let $\vec{x} = \begin{bmatrix} 6 \\ -2 \\ 1 \end{bmatrix}$. What are $[\vec{x}]_{\mathcal{E}}$ and $[\vec{x}]_{\mathcal{B}}$?

Solution

We have

$$\vec{x} = \begin{bmatrix} 6 \\ -2 \\ 1 \end{bmatrix} = 6 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

and therefore

$$[\vec{x}]_{\mathcal{E}} = \begin{bmatrix} 6 \\ -2 \\ 1 \end{bmatrix}$$

Next, we have

$$\vec{x} = \begin{bmatrix} 6 \\ -2 \\ 1 \end{bmatrix} = -2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + 6 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

and therefore

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 1 \\ 6 \end{bmatrix}$$

Notice how $[\vec{x}]_{\mathcal{E}}$ and $[\vec{x}]_{\mathcal{B}}$ differ from each other, due to the order of the basis vectors in $\mathcal{E}$ and $\mathcal{B}$.

Exercise 1

A question appears in Mobius

If we know a basis of $\mathbb{R}^n$, we can determine which vector corresponds to a given coordinate vector. Let's see an example of this.

Example 2

Let $\mathcal{C} = \left\{ \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} \right\}$ be a basis of $\mathbb{R}^3$. If $[\vec{x}]_{\mathcal{C}} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$, what is $\vec{x}$ in standard coordinates?

Solution

The coordinates in $[\vec{x}]_{\mathcal{C}}$ tell us the scalars that multiply each of the basis vectors in $\mathcal{C}$. Thus, $[\vec{x}]_{\mathcal{C}}$ represents the vector

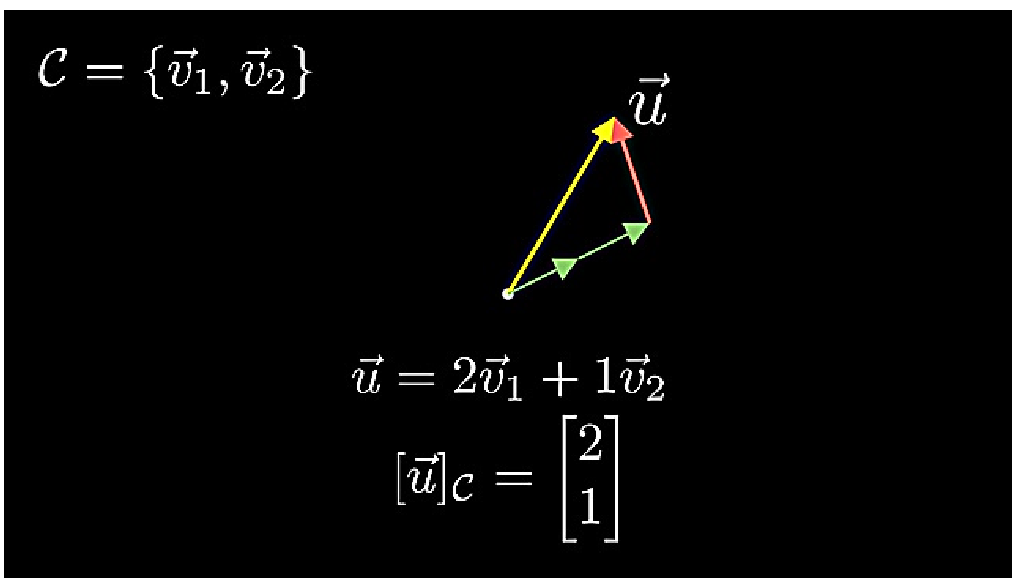
$$3 \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} - 1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 2 \end{bmatrix}$$

Exercise 2

A question appears in Mobius

Consider the standard basis $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ and the vector $\vec{u}_{\mathcal{B}} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$, then $\vec{u} = 3\vec{e}_1 + 5\vec{e}_2$. In contrast, consider the non-standard basis $\mathcal{C} = \{\vec{v}_1, \vec{v}_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \end{bmatrix} \right\}$. Then $\vec{u} = 2\vec{v}_1 + 1\vec{v}_2$ and so $\vec{u}_{\mathcal{C}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. This is demonstrated in the following video.

A video appears here.



Making Connections

Create your own basis $\mathcal{B}$ by dragging the vectors $\vec{b}_1$ and $\vec{b}_2$, then drag the vector $\vec{x}$ to see the $\mathcal{B}$-coordinates of $\vec{x}$.

External resource: https://www.geogebra.org/material/iframe/id/g68dvydw/

4.5 - Change of Basis Matrix

Change of Basis Matrix

Introduction

Remember that our goal is to determine the coordinate vector when using a non-standard basis. Another way to think about this is that we are trying to change coordinates from the standard basis to some other basis. In fact, we can generalize our question: how do we change coordinates from one basis to another? Our general technique will be to find a special matrix, which we will call the **change of basis matrix**.

Let's restate what we're trying to do a bit more formally: given a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ and a basis $\mathcal{C} = \{\vec{w}_1, \ldots, \vec{w}_n\}$ for $\mathbb{R}^n$, we want to take a vector which is expressed in $\mathcal{C}$-coordinates and change it into $\mathcal{B}$-coordinates.

Definition
Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ and $\mathcal{C} = \{\vec{w}_1, \ldots, \vec{w}_n\}$ both be bases for $\mathbb{R}^n$. The matrix ${}_{\mathcal{B}}P_{\mathcal{C}} = \begin{bmatrix} [\vec{w}_1]_{\mathcal{B}} & \cdots & [\vec{w}_n]_{\mathcal{B}} \end{bmatrix}$ is called the **change of basis matrix (or change of coordinates matrix)** from $\mathcal{C}$-coordinates to $\mathcal{B}$-coordinates and satisfies

$$[\vec{x}]_{\mathcal{B}} = {}_{\mathcal{B}} P_{\mathcal{C}}[\vec{x}]_{\mathcal{C}}$$

Notice that the subscripts of ${}_{\mathcal{B}}P_{\mathcal{C}}$ are in the **opposite** order than what is read: the change of coordinates from $\mathcal{C}$-coordinates to $\mathcal{B}$-coordinates is denoted by ${}_{\mathcal{B}}P_{\mathcal{C}}$

Let's see an example of how the change of basis works.

Example 1

Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix} \right\}$ and $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix} \right\}$ be bases for $\mathbb{R}^3$.

Let $\vec{x} \in \mathbb{R}^3$ be such that $[\vec{x}]_{\mathcal{C}} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$. Find $[\vec{x}]_{\mathcal{B}}$.

Solution

We will find $[\vec{x}]_{\mathcal{B}}$ by using the fact that $[\vec{x}]_{\mathcal{B}} = {}_{\mathcal{B}} P_{\mathcal{C}}[\vec{x}]_{\mathcal{C}}$. To do this, we need to find the change of basis matrix, ${}_{\mathcal{B}}P_{\mathcal{C}}$.

${}_{\mathcal{B}}P_{\mathcal{C}}$ is the matrix whose columns are the $\mathcal{B}$-coordinates of the basis vectors in $\mathcal{C}$.

Let's start by finding these $\mathcal{B}$-coordinates.

To find $\left[\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right]_{\mathcal{B}}$, we equate the first element of C to a linear combination of the elements in B with scalars $t_1, t_2, t_3 \in \mathbb{R}$ and solve for t_1, t_2, t_3:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = t_1 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + t_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + t_3 \begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix}$$

We can write this as the following system of linear equations:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ -1 & 1 & 3 & 1 \end{array}\right]$$

To find $\left[\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}\right]_{\mathcal{B}}$, we equate the second element of C to a linear combination of the elements in B with scalars $r_1, r_2, r_3 \in \mathbb{R}$:

$$\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} = r_1 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + r_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + r_3 \begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix}$$

We can write this as the following system of linear equations:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ -1 & 1 & 3 & -2 \end{array}\right]$$

Finally, to find $\left[\begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix}\right]_{\mathcal{B}}$, we equate the third element of C to a linear combination of the elements in B with scalars $s_1, s_2, s_3 \in \mathbb{R}$:

$$\begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix} = s_1 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + s_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + s_3 \begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix}$$

We can write this as the following system of linear equations:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 4 \\ -1 & 1 & 3 & 0 \end{array}\right]$$

Notice that the coefficient matrix for these three systems is the same. We can therefore form the following augmented system to solve all of the equations simultaneously:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & -1 & 1 & -1 & 4 \\ -1 & 1 & 3 & 1 & -2 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1/2 & -7/4 & 3 \\ 0 & 0 & 1 & 1/2 & 1/4 & -1 \end{array}\right]$$

Thus the change of basis matrix is

$${}_{\mathcal{B}}P_{\mathcal{C}} = \begin{bmatrix} 1 & 1 & 0 \\ 1/2 & -7/4 & 3 \\ 1/2 & 1/4 & -1 \end{bmatrix}$$

To find $[\vec{x}]_{\mathcal{B}}$ for $[\vec{x}]_{\mathcal{C}} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$, we multiply:

$$\begin{aligned} [\vec{x}]_{\mathcal{B}} &= {}_{\mathcal{B}}P_{\mathcal{C}}[\vec{x}]_{\mathcal{C}} \\ &= \begin{bmatrix} 1 & 1 & 0 \\ 1/2 & -7/4 & 3 \\ 1/2 & 1/4 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix} \end{aligned}$$

To recap, we solved this question in two main steps:

1. First, we found the change of basis matrix ${}_{\mathcal{B}}P_{\mathcal{C}}$
   - The first column of ${}_{\mathcal{B}}P_{\mathcal{C}}$ is the solution to the augmented system $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ -1 & 1 & 3 & 0 \end{array}\right]$
   - The second column of ${}_{\mathcal{B}}P_{\mathcal{C}}$ is the solution to the augmented system $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ -1 & 1 & 3 & -2 \end{array}\right]$
   - The third column of ${}_{\mathcal{B}}P_{\mathcal{C}}$ is the solution to the augmented system $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 4 \\ -1 & 1 & 3 & 0 \end{array}\right]$
2. Secondly, we found $[\vec{x}]_{\mathcal{B}}$ by multiplying ${}_{\mathcal{B}}P_{\mathcal{C}}[\vec{x}]_{\mathcal{C}}$.

Furthering your Understanding

Making Connections

Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ be the standard basis for $\mathbb{R}^2$ and let $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -3 \\ -1 \end{bmatrix} \right\}$ be another basis for $\mathbb{R}^2$.

Then the matrix ${}_{\mathcal{B}}P_{\mathcal{C}}$ such that $[\vec{x}]_{\mathcal{B}} = {}_{\mathcal{B}}P_{\mathcal{C}}[\vec{x}]_{\mathcal{C}}$ is ${}_{\mathcal{B}}P_{\mathcal{C}} = \begin{bmatrix} 1 & -3 \\ 3 & -1 \end{bmatrix}$.

Drag the vector $\vec{x}$ to see its coordinates with respect to $\mathcal{B}$ and $\mathcal{C}$. Observe that the equality $[\vec{x}]_{\mathcal{B}} = {}_{\mathcal{B}}P_{\mathcal{C}}[\vec{x}]_{\mathcal{C}}$ always holds, no matter what vector $\vec{x}$ you choose.

External resource: https://www.geogebra.org/material/iframe/id/m4vdc7cq/

Let's see another example, this time in $\mathbb{R}^4$.

Example 2

Let $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ be the standard basis for $\mathbb{R}^4$, and let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 8 \\ -3 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 1 \\ 7 \end{bmatrix} \right\}$ be another basis for $\mathbb{R}^4$.

1. Find the change of basis matrix ${}_{\mathcal{S}}Q_{\mathcal{B}}$, from $\mathcal{B}$-coordinates to $\mathcal{S}$-coordinates.
2. Find the change of basis matrix ${}_{\mathcal{B}}P_{\mathcal{S}}$ from $\mathcal{S}$-coordinates to $\mathcal{B}$-coordinates.

Solution - Part A

To find the change of basis matrix ${}_{\mathcal{S}}Q_{\mathcal{B}}$, we solve the augmented sytem

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -1 & 3 & -1 \\ 0 & 1 & 0 & 0 & 2 & 0 & 2 & 4 \\ 0 & 0 & 1 & 0 & 3 & -1 & 8 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 & -3 & 7 \end{array}\right]$$

Fortunately, this system is already row reduced, hence

$${}_{\mathcal{S}}Q_{\mathcal{B}} = \begin{bmatrix} 1 & -1 & 3 & -1 \\ 2 & 0 & 2 & 4 \\ 3 & -1 & 8 & 1 \\ 1 & 2 & -3 & 7 \end{bmatrix}$$

Solution - Part B

To find the change of basis matrix ${}_{\mathcal{B}}P_{\mathcal{S}}$, we solve the augmented sytem

$$\left[\begin{array}{cccc|cccc} 1 & -1 & 3 & -1 & 1 & 0 & 0 & 0 \\ 2 & 0 & 2 & 4 & 0 & 1 & 0 & 0 \\ 3 & -1 & 8 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -3 & 7 & 0 & 0 & 0 & 1 \end{array}\right]$$

After row reduction, we find

$${}_{\mathcal{B}}P_{\mathcal{S}} = \begin{bmatrix} 17/3 & -19/6 & -1/3 & 8/3 \\ 5/3 & -8/3 & 2/3 & 5/3 \\ -5/3 & 2/3 & 1/3 & -2/3 \\ -2 & 3/2 & 0 & -1 \end{bmatrix}$$

Remark
Notice that it is much easier to change coordinates from a non-standard basis to a standard basis.

In the previous example, the system

$$\left[\begin{array}{cccc|cccc} 1 & -1 & 3 & -1 & 1 & 0 & 0 & 0 \\ 2 & 0 & 2 & 4 & 0 & 1 & 0 & 0 \\ 3 & -1 & 8 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & -3 & 7 & 0 & 0 & 0 & 1 \end{array}\right]$$

might look familiar to you, if you recall the algorithm for finding the inverse of a matrix. The next theorem confims this observation.

Theorem 1
Let $\mathcal{B}$ and $\mathcal{C}$ both be bases for $\mathbb{R}^n$. Let ${}_{\mathcal{B}}P_{\mathcal{C}}$ be the change of basis matrix from $\mathcal{C}$-coordinates to $\mathcal{B}$-coordinates. Then ${}_{\mathcal{B}}P_{\mathcal{C}}$ is invertible and ${}_{\mathcal{B}}P_{\mathcal{C}}^{-1}$ is the change of basis matrix from $\mathcal{B}$-coordinates to $\mathcal{C}$-coordinates.

The theorem tells us that once we have a change of basis matrix from $\mathcal{C}$-coordinates to $\mathcal{B}$-coordinates, the inverse of that matrix is the change of basis matrix from $\mathcal{B}$-coordinates to $\mathcal{C}$-coordinates. Let's see how to use this result in practice.

Example 3

Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ and $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \right\}$.

1. Find the change of basis matrix from $\mathcal{C}$-coordinates to $\mathcal{B}$-coordinates.
2. Find the change of basis matrix from $\mathcal{B}$-coordinates to $\mathcal{C}$-coordinates.

Solution - Part A

We notice that $\mathcal{B}$ is the standard basis of $\mathbb{R}^3$, so we should start by finding the change of basis matrix from $\mathcal{C}$-coordinates to $\mathcal{B}$-coordinates. We are solving the augmented system

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

which is already row reduced, therefore the change of basis matrix from $\mathcal{C}$-coordinates to $\mathcal{B}$-coordinates is

$${}_{\mathcal{B}}P_{\mathcal{C}} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

.

Solution - Part B

To find the change of basis matrix from $\mathcal{B}$-coordinates to $\mathcal{C}$-coordinates, we calculate ${}_{\mathcal{B}}P_{\mathcal{C}}^{-1}$ using the inverse matrix algorithm and find

$${}_{\mathcal{C}}P_{\mathcal{B}} = {}_{\mathcal{B}}P_{\mathcal{C}}^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

.

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

Unit 5

Subspaces of $\mathbb{R}^n$

5.1 - Subspaces of $\mathbb{R}^n$

Subspaces of $\mathbb{R}^n$

Definition of Subspaces of $\mathbb{R}^n$

In this unit, we will study special types of sets in $\mathbb{R}^n$ that have certain properties. These sets, called subspaces, behave particularly nicely with the operations of addition and scalar multiplication.

Definition
A set $\mathbb{S}$ in $\mathbb{R}^n$ is a **subspace** of $\mathbb{R}^n$ if

1. $\mathbb{S}$ is a non-empty subset of $\mathbb{R}^n$.
2. $\mathbb{S}$ is closed under addition: that is, for $\vec{s}_1, \vec{s}_2 \in \mathbb{S}$, $\vec{s}_1 + \vec{s}_2 \in \mathbb{S}$.
3. $\mathbb{S}$ is closed under scalar multiplication: that is, for $\vec{s} \in \mathbb{S}$ and $c \in \mathbb{R}$, $c \cdot \vec{s} \in \mathbb{S}$.

Given a subset S of $\mathbb{R}^n$, we are interested in determining whether S is a subspace of $\mathbb{R}^n$ or not. The definition above suggests a natural procedure for this, namely to verify that

1. S is non-empty;
2. S is closed under addition; and
3. S is closed under scalar multiplication.

We refer to this procedure as the **subspace test**.

Let's see how we can apply the **subspace test** to the smallest (non-empty) and the biggest subsets of $\mathbb{R}^n$; namely, to $\{\vec{0}\}$ and to $\mathbb{R}^n$ itself.

For the case where $\mathbb{S} = \{\vec{0}\}$, $\mathbb{S}$ is a subspace of $\mathbb{R}^n$, since:

1. $\{\vec{0}\} \subseteq \mathbb{R}^n$ and $\{\vec{0}\}$ is non-empty;
2. $\vec{0} + \vec{0} = \vec{0}$ (closed under addition), and

3. $c\vec{0} = \vec{0}$ for all $c \in \mathbb{R}$ (closed under multiplication).

For the case where $\mathbb{S} = \mathbb{R}^n$, $\mathbb{S}$ is a subspace of itself, since:

1. $\mathbb{R}^n \subseteq \mathbb{R}^n$, $\mathbb{R}^n$ is non-empty;
2. $\mathbb{R}^n$ is closed under addition; and
3. $\mathbb{R}^n$ is closed under scalar multiplication.

Let's look at some more examples of using the **subspace test**.

Example 1

Is the set $\mathbb{S} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ a subspace of $\mathbb{R}^2$?

Solution

The set $\mathbb{S} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ is **not** a subspace of $\mathbb{R}^2$ since, although it is a non-empty subset of $\mathbb{R}^2$, it is neither closed under addition:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \notin \mathbb{S}$$

nor closed under scalar multiplication:

$$2 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \notin \mathbb{S}$$

Remark
Notice that the set S in the previous example does not contain the zero vector. This already indicates that S is not a subspace: since $cs \in S$ for every $c \in \mathbb{R}$ and every $s \in S$, we can take $c = 0$ which gives $0 \cdot s = \vec{0}$. This means that every subspace must contain the zero vector.

Example 2

Is the set $S = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mid x_1 x_2 = 0 \right\}$ a subspace of $\mathbb{R}^2$?

Solution

S is **not** a subspace of $\mathbb{R}^2$ since it is not closed under addition.

For example, $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \in S$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix} \in S$ since they both satisfy $x_1 x_2 = 0$.

However,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin S$$

since $(1)(1) = 1 \neq 0$.

Example 3

Is the set $S = \emptyset$ a subspace of $\mathbb{R}^3$?

Solution

The empty set is **not** a subspace of $\mathbb{R}^3$ since it it empty. The first part of the subspace test tells us that any subspace must be non-empty.

Example 4

Is the set $\mathbb{S} = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1 + x_2 = 0, x_2 - x_3 = 0 \right\}$ a subspace of $\mathbb{R}^3$?

Solution

We apply the subspace test:

1. By definition, $\mathbb{S} \subseteq \mathbb{R}^3$ and $\mathbb{S}$ contains the zero vector which satisfies $0 + 0 = 0$ and $0 - 0 = 0$. Thus $\mathbb{S}$ is a nonempty subset of $\mathbb{R}^3$.

2. Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in \mathbb{S}$. Since $\vec{x} \in \mathbb{S}$, $x_1 + x_2 = 0$ and $x_2 - x_3 = 0$; since $\vec{y} \in \mathbb{S}$, $y_1 + y_2 = 0$ and $y_2 - y_3 = 0$. Then,
$$\vec{x} + \vec{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix}$$
and we have
$$(x_1 + y_1) + (x_2 + y_2) = (x_1 + x_2) + (y_1 + y_2) = 0 + 0 = 0$$
and
$$(x_2 + y_2) - (x_3 + y_3) = (x_2 - x_3) + (y_2 - y_3) = 0 + 0 = 0$$
Therefore $\vec{x} + \vec{y} \in \mathbb{S}$ so $\mathbb{S}$ is closed under addition.

3. Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{S}$ and let $k \in \mathbb{R}$. Then,
$$k \cdot \vec{x} = \begin{bmatrix} kx_1 \\ kx_2 \\ kx_3 \end{bmatrix}$$
and we have
$$kx_1 + kx_2 = k(x_1 + x_2) = k(0) = 0$$
and
$$kx_2 - kx_3 = k(x_2 - x_3) = k(0) = 0$$
for any scalar $k \in \mathbb{R}$, so $\mathbb{S}$ is closed under scalar multiplication.

Hence $\mathbb{S}$ is a subspace of $\mathbb{R}^3$ by the subspace test.

Remark

To show that $\mathbb{S}$ was non-empty in the previous example, we showed that the zero vector $\vec{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ was an element of $\mathbb{S}$. Showing that an arbitrary set $\mathbb{S}$ is non-empty is typically done by showing that the zero vector is an element of $\mathbb{S}$ since the zero vector is the easiest vector to work with.

Exercise 1

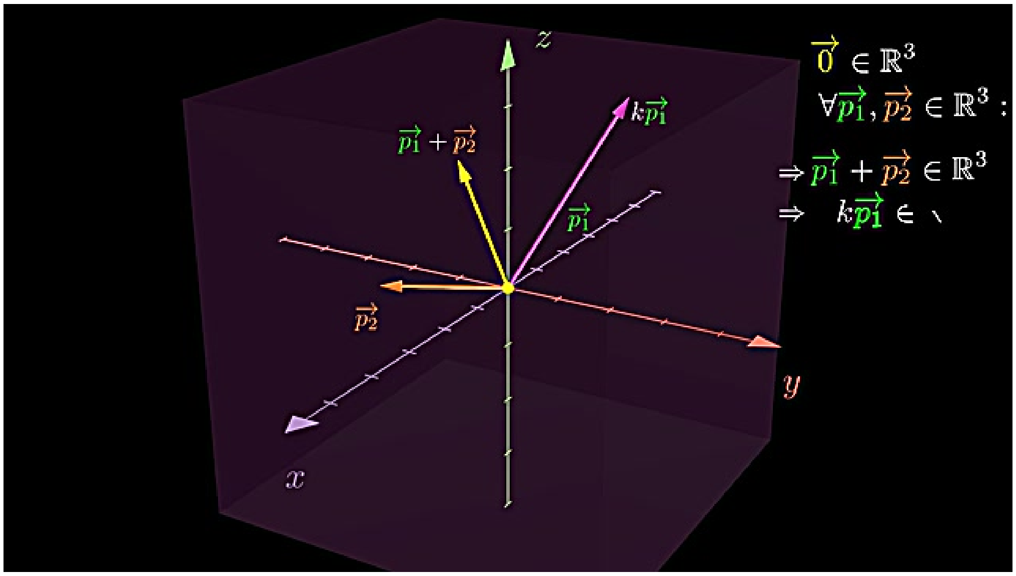
A question appears in Mobius

Exercise 2

A question appears in Mobius

Watch the following video, which illustrates the possible subspaces of $\mathbb{R}^3$: the point at the origin $\{\vec{0}\}$, lines through the origin, planes through the origin, and $\mathbb{R}^3$ itself. Note that this video has no sound.

A video appears here.



Span and Subspaces of $\mathbb{R}^n$

The span of a set of vectors in $\mathbb{R}^n$ has the interesting property that it is always a subspace of $\mathbb{R}^n$. We will see an example of this, and you will have the chance to prove the general result for yourself afterwards.

Example 5

Show that $S = \text{Span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}-2\\0\\3\end{bmatrix}\right\}$ is a subspace of $\mathbb{R}^3$.

Solution

We prove this using the subspace test:

1. By definition, S is the set of all linear combinations of $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}-2\\0\\3\end{bmatrix}$. This is a subset of $\mathbb{R}^3$ and furthermore, $\begin{bmatrix}0\\0\\0\end{bmatrix} = 0\begin{bmatrix}1\\1\\1\end{bmatrix} + 0\begin{bmatrix}-2\\0\\3\end{bmatrix}$ so $\vec{0} \in S$ which means that S is non-empty.

2. Let $\vec{x}, \vec{y} \in \mathbb{S}$. Then $\vec{x}$ and $\vec{y}$ are both linear combinations of the vectors $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}-2\\0\\3\end{bmatrix}$. Thus, they have the form
$$\vec{x} = c_1\begin{bmatrix}1\\1\\1\end{bmatrix} + c_2\begin{bmatrix}-2\\0\\3\end{bmatrix} \quad \text{and} \quad \vec{y} = d_1\begin{bmatrix}1\\1\\1\end{bmatrix} + d_2\begin{bmatrix}-2\\0\\3\end{bmatrix}$$

for scalars $c_1, c_2, d_1, d_2 \in \mathbb{R}$.
Let's verify that their sum, $\vec{x}+\vec{y}$ is also in S:

$$\begin{aligned}\vec{x}+\vec{y} &= c_1\begin{bmatrix}1\\1\\1\end{bmatrix}+c_2\begin{bmatrix}-2\\0\\3\end{bmatrix}+d_1\begin{bmatrix}1\\1\\1\end{bmatrix}+d_2\begin{bmatrix}-2\\0\\3\end{bmatrix}\\ &= (c_1+d_1)\begin{bmatrix}1\\1\\1\end{bmatrix}+(c_2+d_2)\begin{bmatrix}-2\\0\\3\end{bmatrix}\end{aligned}$$

Notice that the sum of $\vec{x}+\vec{y}$ still gives a linear combination of $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}-2\\0\\3\end{bmatrix}$, so $\mathbb{S}$ is closed under addition.

3. Let $\vec{x} \in \mathbb{S}$ and $t \in \mathbb{R}$. Then

$$t\vec{x} = t\left(c_1\begin{bmatrix}1\\1\\1\end{bmatrix}+c_2\begin{bmatrix}-2\\0\\3\end{bmatrix}\right) = (tc_1)\begin{bmatrix}1\\1\\1\end{bmatrix}+(tc_2)\begin{bmatrix}-2\\0\\3\end{bmatrix}$$

The scalar product of $t\vec{x}$ still gives a linear combination of $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}-2\\0\\3\end{bmatrix}$, so $\mathbb{S}$ is closed under scalar multiplication.

By the subspace test, $S = \text{Span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}-2\\0\\3\end{bmatrix}\right\}$ is a subspace of $\mathbb{R}^3$.

Exercise 3

Let $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$. Show that $\mathbb{S} = \text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\}$ is a subspace of $\mathbb{R}^n$.

A question appears in Mobius

Exercise 4

A question appears in Mobius

5.2 - Basis and Dimension

Basis of a Subspace of $\mathbb{R}^n$

Finding a Basis of a Subspace of $\mathbb{R}^n$

In the previous lesson, we encountered subspaces of $\mathbb{R}^n$; the last exercise showed that a non-empty set of vectors in $\mathbb{R}^n$ spans a subspace of $\mathbb{R}^n$. In fact, every subspace of $\mathbb{R}^n$ (except for the trivial subspace) is spanned by a set of vectors belonging to that subspace.

Example 1

Find a basis for the subspace $\mathbb{S} = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1 + x_2 = 0 \quad \text{and} \quad x_2 - x_3 = 0 \right\}$ of $\mathbb{R}^3$.

Solution

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{S}$. Then $x_1 + x_2 = 0$ and $x_2 - x_3 = 0$.

We can rearrange these equations to find $x_1 = -x_2$ and $x_3 = x_2$.

Substituting these values into $\vec{x}$, we have

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad x_2 \in \mathbb{R}$$

Thus, $\mathbb{S} \subseteq \text{Span}\left\{\begin{bmatrix}-1\\1\\1\end{bmatrix}\right\}$ and therefore $B = \left\{\begin{bmatrix}-1\\1\\1\end{bmatrix}\right\}$ is a spanning set of $\mathbb{S}$.

And, since B consists of a single non-zero vector, B is linearly independent and hence a basis for $\mathbb{S}$.

Note that now that we have a basis for $\mathbb{S}$, we recognize that $\mathbb{S}$ is a line passing through the origin.

Let's have a closer look at lines passing through the origin, and what this means.

Making Connections

Move the line by dragging points A and B. When the line passes through the origin, it becomes clear that the vectors $\vec{u}$, $\vec{v}$, and $\vec{u} + \vec{v}$ lie on this line.

This serves as a visual demonstration of the fact that a line passing through the origin is always a subspace. When a line does not pass through the origin, it is no longer a subspace, since it is not closed under addition (i.e. the vector u+v does not lie on the line.)

External resource: https://www.geogebra.org/material/iframe/id/tx4hcevt/

Algorithm
To find a basis for a subspace $\mathbb{S}$ of $\mathbb{R}^n$, we need to find a subset $\{\vec{v}_1, \ldots, \vec{v}_k\}$ of $\mathbb{S}$ which:

- spans $\mathbb{S}$ and
- is linearly independent.

We do this as follows:

1. Choose an arbitrary $\vec{x} \in \mathbb{S}$ and try to "decompose" $\vec{x}$ as a linear combination of some $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{S}$. This shows that $\mathbb{S} \subseteq \text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\}$.
2. Verify that the set $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent using your favourite method.

Remark
For part 1 in the above algorithm, to show that $\mathbb{S} = \text{Span}\{\vec{v}_1, \ldots, \vec{v}_n\}$, we should also show that $\text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\} \subseteq \mathbb{S}$. However, we don't need to do this explicitly, since $\mathbb{S}$ is a subspace and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a subset of $\mathbb{S}$.

Example 2

Given the subspace $\mathbb{S} = \left\{\begin{bmatrix}a-b\\b-c\\c-a\end{bmatrix} \in \mathbb{R}^3 \mid a, b, c, \in \mathbb{R}\right\}$ of $\mathbb{R}^3$, find a basis for $\mathbb{S}$.

Solution

Let $\vec{x} \in \mathbb{S}$. Then,

$$\vec{x} = \begin{bmatrix} a-b \\ b-c \\ c-a \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

for $a, b, c \in \mathbb{R}$. Thus,

$$\mathbb{S} = \operatorname{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$$

So $B = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ is a spanning set for $\mathbb{S}$.

Note that

$$\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = - \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

so we can remove this vector from the set B. Therefore

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

is a spanning set for $\mathbb{S}$ and since neither vector in B is a scalar multiple of the other, B is linearly independent and hence a basis for $\mathbb{S}$.

Now that we have a basis for $\mathbb{S}$, we recognize that $\mathbb{S}$ is a plane passing through the origin.

Orthogonal and Orthonormal Sets and Bases

Orthogonal Sets and Bases

We commonly work with a type of set called an orthogonal set.

Definition
A set $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an **orthogonal set** if $\vec{v}_i \cdot \vec{v}_j = 0$ for $i \neq j$. In other words, the dot product of each pair of **distinct** vectors in the set is equal to zero.

Remark
An orthogonal set may contain the zero vector, which would make the set linearly dependent.

Making Connections

Create a non-zero vector $\vec{v}$ that is orthogonal to the given vector $\vec{u}$, by dragging the tip of the vector $\vec{v}$ to an appropriate location in the xy-plane. The feedback indicates whether the vector you created is non-zero and is orthogonal to $\vec{u}$. You can refresh the app to get a different vector $\vec{u}$.

External resource: https://www.geogebra.org/material/iframe/id/ftxj9u7t/

Example 1

- The set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ in $\mathbb{R}^3$ is orthogonal since

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 1(0) + 0(1) + 0(0) = 0$$

- The set $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is orthogonal since

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = 1(1) + 1(-2) + 1(1) = 0$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 1(-1) + 1(0) + 1(1) = 0$$

$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 1(-1) + (-2)(0) + 1(1) = 0$$

The following theorem gives a useful relationship between orthogonal sets of non-zero vectors and linearly independent sets:

> Theorem 1
> If $\{\vec{v}_1, \ldots, \vec{v}_k\} \subseteq \mathbb{R}^n$ is an orthogonal set of non-zero vectors, then $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent.

Proof

For $c_1, \ldots, c_k$ consider the equation $c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$.

For $1 \leq i \leq k$, we have

$$\vec{v}_i \cdot \underbrace{(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k)}_{\vec{0}} = \vec{v}_i \cdot \vec{0}$$

Working on the left-hand side of the equation, we can expand using $\vec{v}_i \cdot \vec{v}_j = 0$ for $i \neq j$ to get

$$c_1 \underbrace{\vec{v}_1 \cdot \vec{v}_i}_{0} + \cdots + c_i \underbrace{\vec{v}_i \cdot \vec{v}_i}_{\|\vec{v}_i\|^2} + \cdots + c_k \underbrace{\vec{v}_k \cdot \vec{v}_i}_{0} = \underbrace{\vec{v}_i \cdot \vec{0}}_{0} \quad \text{since } \vec{v}_j \cdot \vec{v}_i = 0 \text{ for } i \neq j$$

that is,

$$c_i\|\vec{v}_i\|^2 = 0$$

Since $\vec{v}_i \neq \vec{0}$, it follows that $\|\vec{v}_i\|^2 \neq 0$ so it must be that $c_i = 0$ for $1 \leq i \leq k$, hence $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent.

□

Exercise 1

This question has four parts, each with many possible solutions. Consider how you would answer each one, and then click and reveal to see a sample solution.

State a set of three vectors in $\mathbb{R}^3$ which are:

1. Orthogonal and linearly independent.
2. Orthogonal and linearly dependent.
3. Not orthogonal and linearly independent.
4. Not orthogonal and linearly dependent.

A question appears in Mobius

Definition
If an orthogonal set B is a basis for a subspace $\mathbb{S}$ of $\mathbb{R}^n$, then B is an **orthogonal basis** for $\mathbb{S}$.

Example 2

Is the set $B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$ an orthogonal basis for $\mathbb{R}^3$?

Solution

You can check for yourself that B is a basis for $\mathbb{R}^3$. Furthermore, B is orthogonal since

$$\begin{bmatrix}1\\1\\1\end{bmatrix} \cdot \begin{bmatrix}-2\\1\\1\end{bmatrix} = 1(-2) + 1(1) + 1(1) = 0$$
$$\begin{bmatrix}1\\1\\1\end{bmatrix} \cdot \begin{bmatrix}0\\1\\-1\end{bmatrix} = 1(0) + 1(1) + 1(-1) = 0$$
$$\begin{bmatrix}-2\\1\\1\end{bmatrix} \cdot \begin{bmatrix}0\\1\\-1\end{bmatrix} = (-2)(0) + 1(1) + 1(-1) = 0$$

Exercise 2

A question appears in Mobius

Orthogonal bases have the following useful property: if $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal basis for a subspace $\mathbb{S}$ of $\mathbb{R}^n$ and $\vec{x} \in \mathbb{S}$, then there are unique scalars $c_1, \ldots, c_k \in \mathbb{R}$ such that

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

For any $1 \leq i \leq k$, we have

$$\vec{v}_i \cdot \vec{x} = \vec{v}_i(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k) = c_i\|\vec{v}_i\|^2$$

which we can rearrange to get

$$c_i = \frac{\vec{x} \cdot \vec{v}_i}{\|\vec{v}_i\|^2}$$

substituting the c_i back into the equation for $\vec{x}$, we have

$$\vec{x} = \frac{\vec{x} \cdot \vec{v}_1}{\|\vec{v}_1\|^2}\vec{v}_1 + \cdots + \frac{\vec{x} \cdot \vec{v}_k}{\|\vec{v}_k\|^2}\vec{v}_k$$

Example 3

Let $\vec{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $B = \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ -2 \end{bmatrix} \right\}$ be an orthogonal basis for $\mathbb{R}^2$. Write $\vec{x}$ as a linear combination of the vectors in B.

Solution

From the property above, we know that we can write $\vec{x}$ uniquely as

$$\vec{x} = c_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

where

$$c_1 = \frac{\begin{bmatrix} -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix}}{\left\| \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\|^2} = \frac{-1+3}{1+9} = \frac{1}{5}$$
$$c_2 = \frac{\begin{bmatrix} -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ -2 \end{bmatrix}}{\left\| \begin{bmatrix} 6 \\ -2 \end{bmatrix} \right\|^2} = \frac{-6-2}{36+4} = \frac{-1}{5}$$

So we have

$$\vec{x} = \frac{1}{5} \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \frac{1}{5} \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

Exercise 3

A question appears in Mobius

Orthonormal Sets and Bases

Definition
An orthogonal set $\{\vec{v}_1, \ldots, \vec{v}_k\} \subseteq \mathbb{R}^n$ is an **orthonormal set** if $\|\vec{v}_i\| = 1$ for $1 \leq i \leq k$. If an orthonormal set B is a basis for a subspace $\mathbb{S}$ of $\mathbb{R}^n$, then B is an **orthonormal basis** for $\mathbb{S}$.

Example 4

- The set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ in $\mathbb{R}^3$ is orthonormal since it is orthogonal:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 1(0) + 0(1) + 0(0) = 0$$

  and the norm of each vector is 1:

$$\left\| \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\| = \sqrt{1^2 + 0^2 + 0^2} = 1 \quad \text{and} \quad \left\| \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\| = \sqrt{0^2 + 1^2 + 0^2} = 1$$

- The set $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is not orthonormal. It is orthogonal since

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = 1(1) + 1(-2) + 1(1) = 0$$
$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 1(-1) + 1(0) + 1(1) = 0$$
$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 1(-1) + (-2)(0) + 1(1) = 0$$

  but it is not orthonormal since at least one of the norms is not equal to one. For instance,

$$\left\| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3} \neq 1$$

A question appears in Mobius

A question appears in Mobius

Example 5

Is the set $B = \left\{ \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \right\}$ an orthonormal basis for $\mathbb{R}^3$?

Solution

You can check for yourself that B is indeed a basis for $\mathbb{R}^3$.

B is orthogonal since

$$\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} = (1/\sqrt{3})(1/\sqrt{6}) + (1/\sqrt{3})(-2/\sqrt{6}) + (1/\sqrt{3})(1/\sqrt{6}) = 0$$

$$\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \cdot \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} = (1/\sqrt{3})(-1/\sqrt{2}) + (1/\sqrt{3})(0) + (1/\sqrt{3})(1/\sqrt{2}) = 0$$

$$\begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \cdot \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} = (1/\sqrt{6})(-1/\sqrt{2}) + (-2/\sqrt{6})(0) + (1/\sqrt{6})(1/\sqrt{2}) = 0$$

and B is orthonormal since

$$\left\| \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \right\| = 1/3 + 1/3 + 1/3 = 1$$

$$\left\| \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \right\| = 1/6 + 4/6 + 1/6 = 1$$

$$\left\| \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \right\| = 1/2 + 0 + 1/2 = 1$$

Remark
The condition $\|\vec{v}_i\| = 1$ excludes the zero vector from any orthonormal set. Since orthonormal sets are orthogonal sets, orthonormal sets are linearly independent.

Given an orthogonal basis $B = \{\vec{v}_1, \dots, \vec{v}_k\}$ of a subspace $\mathbb{S}$ of $\mathbb{R}^n$, we can obtain an orthonormal basis $C = \{\vec{w}_1, \dots, \vec{w}_n\}$ by setting

$$\vec{w}_i = \frac{1}{\|\vec{v}_i\|}\vec{v}_i \quad \text{for } 1 \le i \le k$$

Orthonormal bases also satisfy the useful property that we previously discussed for orthogonal bases. In fact, the property is further simplified by the fact that the basis vectors are orthonormal: if $B = \{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthonormal basis for a subspace $\mathbb{S}$ of $\mathbb{R}^n$ and $\vec{x} \in \mathbb{S}$, then there are unique scalars $c_1, \dots, c_k \in \mathbb{R}$ such that

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

And the c_i satisfy

$$c_i = \frac{\vec{x} \cdot \vec{v}_i}{\|\vec{v}_i\|^2} = \frac{\vec{x} \cdot \vec{v}}{1} = \vec{x} \cdot \vec{v}_i$$

so we have

$$\vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k$$

Example 6

Let $\vec{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $B = \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ -2 \end{bmatrix} \right\}$ be an orthogonal basis for $\mathbb{R}^3$.

1. Make B into an orthonormal basis C and
2. Write $\vec{x}$ as a linear combination of the vectors in C.

Solution - Part A

In order to make B into an orthonormal basis C, we need to scale each vector in B by its norm. Letting $\vec{v}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$, we set

$$\begin{aligned}
\vec{w}_1 &= \frac{1}{\left\| \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\|} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} \\
\vec{w}_2 &= \frac{1}{\left\| \begin{bmatrix} 6 \\ -2 \end{bmatrix} \right\|} \begin{bmatrix} 6 \\ -2 \end{bmatrix} = \frac{1}{2\sqrt{10}} \begin{bmatrix} 6 \\ -2 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{10} \end{bmatrix}
\end{aligned}$$

Hence

$$C = \left\{ \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix}, \begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{10} \end{bmatrix} \right\}$$

is an orthonormal basis for $\mathbb{R}^2$.

Solution - Part B

Since

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} = \frac{2}{\sqrt{10}} \quad \text{and} \quad \begin{bmatrix} -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{10} \end{bmatrix} = \frac{-4}{\sqrt{10}}$$

we have

$$\vec{x} = \vec{x} \cdot \vec{w}_1 + \vec{x} \cdot \vec{w}_2 = \frac{2}{\sqrt{10}} \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} - \frac{4}{\sqrt{10}} \begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{10} \end{bmatrix}$$

Exercise 5

A question appears in Mobius

Dimension

Finding the Dimension of a Subspace of $\mathbb{R}^n$

In this section, we wil define the dimension of a subspace of $\mathbb{R}^n$. Before we get to the main definition, we will make a few observations.

Let's start by supposing that $\mathbb{S}$ is a subspace of $\mathbb{R}^n$ and that $B = \{\vec{v}_1, \vec{v}_2\}$ is a basis for $\mathbb{S}$. If $C = \{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ is a set of vectors in $\mathbb{S}$, then C is linearly dependent. Let's see why.

Since B is a basis for $\mathbb{S}$, we can write

$$\begin{aligned}\vec{w}_1 &= a_1\vec{v}_1 + a_2\vec{v}_2\\ \vec{w}_2 &= b_1\vec{v}_1 + b_2\vec{v}_2\\ \vec{w}_3 &= c_1\vec{v}_1 + c_2\vec{v}_2\end{aligned}$$

for $a_1, a_2, b_1, b_2, c_1, c_2 \in \mathbb{R}$.

Now, for $t_1, t_2, t_3 \in \mathbb{R}$, consider

$$\begin{aligned}\vec{0} &= t_1\vec{w}_1 + t_2\vec{w}_2 + t_3\vec{w}_3\\ &= t_1(a_1\vec{v}_1 + a_2\vec{v}_2) + t_2(b_1\vec{v}_1 + b_2\vec{v}_2) + t_3(c_1\vec{v}_1 + c_2\vec{v}_2)\\ &= (a_1t_1 + b_1t_2 + c_1t_3)\vec{v}_1 + (a_2t_1 + b_2t_2 + c_2t_3)\vec{v}_2\end{aligned}$$

Since B is linearly independent, we must have

$$\begin{aligned}a_1t_1 + b_1t_2 + c_1t_3 &= 0\\ a_2t_1 + b_2t_2 + c_2t_3 &= 0\end{aligned}$$

which is a homogeneous system (see the lesson on homogeneous systems) with two equations and three unknowns. The system is underdetermined and therefore has non-trivial solutions, which means that $C = \{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ is linearly dependent.

This first observation can be generalized with the following theorem.

Theorem 2
Let $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ be a basis for a subspace $\mathbb{S}$ of $\mathbb{R}^n$. If $C = \{\vec{w}_1, \ldots, \vec{w}_\ell\}$ is a set in $\mathbb{S}$ with $\ell > k$, then C is linearly dependent.

We won't provide a proof of the theorem here, but we will use it to prove our main result, which is the following theorem.

Theorem 3
If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ and $C = \{\vec{w}_1, \ldots, \vec{w}_\ell\}$ are both bases for a subspace $\mathbb{S}$ of $\mathbb{R}^n$, then $k = \ell$.

Proof

Since B is a basis for $\mathbb{S}$ and C is linearly independent, we have $\ell \leq k$ by the theorem above. Also, since C is a basis for $\mathbb{S}$ and B is linearly independent, we have $k \leq \ell$ by the same theorem. Therefore $k = \ell$.

$\square$

We can now make the following definition:

Definition
If $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is a basis for a subspace $\mathbb{S}$ of $\mathbb{R}^n$, the **dimension** of $\mathbb{S}$ is k and we write $\dim(\mathbb{S}) = k$. If $\mathbb{S} = \{\vec{0}\}$, then $\dim(\mathbb{S}) = 0$ since $\emptyset$ is a basis for $\mathbb{S}$.

Example 1

Find the dimension of $\mathbb{R}^n$.

Solution

Recall that $\mathbb{R}^n$ has the standard basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ since this basis has n elements. Therefore the dimension of $\mathbb{R}^n$ is equal to n, i.e., $\dim(\mathbb{R}^n) = n$. This means that **any** basis of $\mathbb{R}^n$ has exactly n vectors.

Example 2

Find the dimension of the subspace $\mathbb{S} = \left\{ \begin{bmatrix} a-b \\ b-c \\ c-a \end{bmatrix} \in \mathbb{R}^3 \mid a, b, c, \in \mathbb{R} \right\}$ of $\mathbb{R}^3$, given that one possible basis for S is $B = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$.

Solution

The dimension of S is equal to 2, i.e., $\dim(\mathbb{S}) = 2$. This means that **any** basis of $\mathbb{S}$ has exactly 2 vectors.

Example 3

Find the dimension of the subspace $\mathbb{S} = \left\{\vec{x} \in \mathbb{R}^4 \mid \vec{x} \cdot \vec{n} = 0\right\}$, where $\vec{n} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}$.

Solution

To find the dimension of S, we first compute its basis:

$$\mathbb{S} = \left\{\vec{x} \in \mathbb{R}^4 \mid x_1 - x_3 + 2x_4 = 0\right\}$$

Rearranging, we can rewrite $x_1 - x_3 + 2x_4 = 0$ as $x_1 = x_3 - 2x_4$. We have:

$$S = \left\{ \begin{bmatrix} x_3 - 2x_4 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \mid x_2, x_3, x_4 \in \mathbb{R} \right\}$$

$$= \left\{ x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \mid x_2, x_3, x_4 \in \mathbb{R} \right\}$$
$$= \text{Span} \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Since the set $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is linearly independent and spans $\mathbb{S}$, it is a basis of S. Since this set contains 3 elements, we conclude that $\dim(\mathbb{S}) = 3$.

We will now state a very useful and important theorem:

Theorem 4
If $\mathbb{S}$ is a k-dimensional subspace of $\mathbb{R}^n$ with $k > 0$, then

1. A set of more than k vectors in $\mathbb{S}$ must be linearly dependent.
2. A set of fewer than k vectors in $\mathbb{S}$ cannot span $\mathbb{S}$.
3. A set of k vectors in $\mathbb{S}$ spans $\mathbb{S}$ if and only if it is linearly independent.

Note that we cannot use this theorem unless we know $\dim(\mathbb{S})$.

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

5.3 - Fundamental Subspaces of a Matrix

Fundamental Subspaces of a Matrix

So far, we have discussed subspaces of $\mathbb{R}^n$. In this section, we will see three subspaces associated to an $m \times n$ matrix A whose entries are real numbers. These subspaces are collectively called the **fundamental subspaces** of a matrix.

To help with notation, we denote the set of matrices $m \times n$ matrices with real entries by $M_{m\times n}(\mathbb{R})$, so $A \in M_{m\times n}(\mathbb{R})$means that A is an $m \times n$ matrix with real entries.

The Nullspace of A

Definition
Let $A \in M_{m\times n}(\mathbb{R})$. The **nullspace** of A is

$$\text{Null}(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}_{\mathbb{R}^m}\}$$

The nullspace of A is the set of all vectors $\vec{x} \in \mathbb{R}^n$ for which the product $A\vec{x}$ gives the zero vector in $\mathbb{R}^m$. We have encountered this subspace before as the solution space of the homogeneous system $A\vec{x} = \vec{0}$. Notice that Null(A) is a subspace of $\mathbb{R}^n$.

Example 1

Let $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & -1 \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$.

1. Find a vector that is in the nullspace of A.
2. Find a vector that is **not** in the nullspace of A

Solution - Part A

The vector $\vec{x} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$ is in the nullspace of A since

$$A\vec{x} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1(2) + 0(-2) + (-1)(2) \\ 2(2) + 1(-2) + (-1)(2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solution - Part B

The vector $\vec{y} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ is not in the nullspace of A since

$$A\vec{y} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1(1) + 0(2) + (-1)(1) \\ 2(1) + 1(2) + (-1)(1) \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Exercise 1

A question appears in Mobius

The Column Space of A

Definition
Let $A \in M_{m\times n}(\mathbb{R})$ be the matrix with columns given by the vectors $\{\vec{a}_1, \ldots, \vec{a}_n\}$. The **column space** of A is

$$\mathrm{Col}(A) = \{A\vec{x} \mid \vec{x} \in \mathbb{R}^n\} = \mathrm{Span}\{\vec{a}_1, \ldots, \vec{a}_n\}$$

The columns space of A contains all of the vectors that can be obtained by taking linear combinations of the columns of the matrix A. Since the column space of A is spanned by the columns of A, it is a subspace of $\mathbb{R}^m$.

Example 2

Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} \in M_{3\times 2}(\mathbb{R})$.

1. Find a vector that is in the column space of A.
2. Find a vector that is **not** in the column space of A.

Solution - Part A

The vector $\vec{x} = \begin{bmatrix} 5 \\ 2 \\ -3 \end{bmatrix}$ is in the column space of A since

$$\begin{bmatrix} 5 \\ 2 \\ -3 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}.$$

Solution - Part B

The vector $\vec{y} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ is not in the column space of A since $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ cannot be written as a linear combination of $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$.

Remark
From our previous work, we know that the system $A\vec{x} = \vec{b}$ is consistent if and only if $\vec{b}$ is a linear combination of the columns of A. Now we can say that $A\vec{x} = \vec{b}$ is consistent if and only if $\vec{b} \in \text{Col}(A)$.

Exercise 2

A question appears in Mobius

The Row Space of A

Definition
Let $A \in M_{m\times n}(\mathbb{R})$ be the matrix with rows given by the vectors $\{\vec{r}_1^{\,T}, \ldots, \vec{r}_m^{\,T}\}$. The **row space** of A is

$$\text{Row}(A) = \{A^T\vec{x} \mid \vec{x} \in \mathbb{R}^m\} = \text{Span}\{\vec{r}_1, \ldots, \vec{r}_m\}$$

The row space of A contains all of the vectors that can be obtained by taking linear combinations of the rows of the matrix A. Since the row space of A is spanned by the rows of A, it is a subspace of $\mathbb{R}^n$.

Example 3

Let $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$.

1. Find a vector that is in the row space of A.

2. Find a vector that is **not** in the row space of A.

Solution - Part A

The vector $\vec{x} = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}$ is in the row space of A since

$$\begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} = 2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Solution - Part B

The vector $\vec{y} = \begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix}$ is not in the row space of A since $\begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix}$ cannot be written as a linear combination of $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$.

Remark
Similar to our previous observation about the column space, $A^T\vec{x} = \vec{b}$ is consistent if and only if $\vec{b} \in \text{Row}(A^T)$ - note the transpose!

Exercise 3

A question appears in Mobius

Finding a Basis for the Fundamental Subspaces

Being subspaces, the fundamental subspaces of a matrix A will each have a basis and dimension. We will now discuss how to find a basis and the dimension of each fundamental subspace. The key will be to take the matrix A into either REF or RREF, depending on the fundamental subspace:

- The **nullspace** of A is the solution set of the homogeneous system $A\vec{x} = \vec{0}$. To find the solution set, and hence the basis vectors for the nullspace, it is necessary to carry A to RREF.
- The **column space** of A is the set spanned by a linearly independent subset of the columns of A. To find the column space, it is enough to carry A into REF and identify the columns with the leading entries. The corresponding columns of A form a basis for the column space of A.
- The **row space** of A is the set spanned by a linearly independent subset of the rows of A. To find the row space, it is enough to carry A into REF and identify the non-zero rows. The set containing the transpose of the non-zero rows (those with leading entries) in the REF forms a basis for the row space of A.

Let us see an example of finding bases and stating the dimensions of fundamental subspaces of a matrix.

A slideshow appears in Mobius.

Slide

Example 4

Let $A = \begin{bmatrix} 1 & 1 & 5 & 1 \\ 1 & 2 & 7 & 2 \\ 2 & 3 & 12 & 3 \end{bmatrix}$. Find a basis for, and state the dimensions of the following subspaces:

1. Null(A)
2. Col(A)
3. Row(A)

Solution

First, we compute the RREF R of A because we will need it to answer all three parts of this question:

$$\begin{bmatrix} 1 & 1 & 5 & 1 \\ 1 & 2 & 7 & 2 \\ 2 & 3 & 12 & 3 \end{bmatrix} \underset{\substack{R_2-R_1 \\ R_3-2R_1}}{\sim} \begin{bmatrix} 1 & 1 & 5 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 \end{bmatrix} \underset{R_3-R_2}{\sim} \begin{bmatrix} 1 & 1 & 5 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \underset{R_1-R_2}{\sim} \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus,

$$R = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Slide

Example 4 - Part A: Finding a Basis of $\text{Null}(A)$

To find a basis for $\text{Null}(A)$, consider the homogeneous system $\left[R \mid \vec{0}\right]$ corresponding to R:

$$\begin{array}{ccccccccc} x_1 & & & + & 3x_3 & & & = & 0 \\ & & x_2 & + & 2x_3 & + & x_4 & = & 0 \end{array}$$

Notice that x_3 and x_4 are free variables. Rearranging these equations gives us:

$$\begin{array}{ccccc} x_1 & = & -3x_3 & & \\ x_2 & = & -2x_3 & -x_4 \end{array}$$

Therefore, the solution space of this system is given by

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3x_3 \\ -2x_3 - x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -3 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \quad x_3, x_4 \in \mathbb{R}$$

We conclude that $\left\{ \begin{bmatrix} -3 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of $\text{Null}(A)$.
Since there are two elements in the basis, $\dim(\text{Null}(A)) = 2$.
Notice that the properties of RREF guarantee that this set is linearly independent.

Slide

Example 4 - Part B: Finding a Basis of $\text{Col}(A)$

To find a basis for $\text{Col}(A)$, we look at the RREF R of A that we found earlier:

$$\underbrace{\begin{bmatrix} 1 & 1 & 5 & 1 \\ 1 & 2 & 7 & 2 \\ 2 & 3 & 12 & 3 \end{bmatrix}}_{\text{A}} \sim \underbrace{\begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{\text{R}}$$

Notice that the first and second columns of R contain the leading ones.
This means that the first and second columns of A constitute a basis of $\text{Col}(A)$.
We conclude that $\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$ is a basis of $\text{Col}(A)$.
Since there are two elements in the basis, $\dim(\text{Col}(A)) = 2$.

Slide

Example 4 - Part C: Finding a Basis of $\mathrm{Row}(A)$

To find a basis for $\mathrm{Row}(A)$, we again look at the RREF R of A:

$$\underset{\text{A}}{\begin{bmatrix}1 & 1 & 5 & 1\\ 1 & 2 & 7 & 2\\ 2 & 3 & 12 & 3\end{bmatrix}} \sim \underset{\text{R}}{\begin{bmatrix}1 & 0 & 3 & 0\\ 0 & 1 & 2 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}}$$

Notice that the first and second rows of R contain the leading ones.
This means that we can take the transposes of these rows of R to form a basis.

We conclude that $\left\{\begin{bmatrix}1\\0\\3\\0\end{bmatrix}, \begin{bmatrix}0\\1\\2\\1\end{bmatrix}\right\}$ is a basis of $\mathrm{Row}(A)$.

Since there are two elements in the basis, $\dim(\mathrm{Row}(A)) = 2$.

Slide

Summary

Let's summarize how to find bases for the fundamental subspaces of A:
First, find the RREF R of A. Then:

- A basis for $\mathrm{Null}(A)$ is given by the solution to the system $\left[R \mid \vec{0}\right]$.
- A basis for $\mathrm{Col}(A)$ is given by the **columns of** A corresponding to the leading ones in the RREF of A.
- A basis for $\mathrm{Row}(A)$ is given by the transpose of the **rows in the RREF of** A that have leading ones.

Remark
It's important to make sure that you read the information from the RREF (or REF) of A correctly for the column space and the row space.

- Basis elements for the column space are the **columns of** A that correspond to columns with leading entries in the RREF of A.
- Basis elements for the row space are the transpose of **non-zero rows of the RREF of** A.

Example 5

Let $A = \begin{bmatrix}1 & 2 & 1 & 3 & 4\\ 3 & 6 & 2 & 6 & 9\\ -2 & -4 & 1 & 1 & -1\end{bmatrix}$. Find a basis for $\mathrm{Null}(A)$, $\mathrm{Col}(A)$, and $\mathrm{Row}(A)$ and state their dimensions.

Solution

Taking A into RREF, we have: $\begin{bmatrix} 1 & 2 & 1 & 3 & 4 \\ 3 & 6 & 2 & 6 & 9 \\ -2 & -4 & 1 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$

The nullspace of A is the solution set to the homogeneous system $A\vec{x} = \vec{0}$. The general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \quad s, t, \in \mathbb{R}$$

so

$$B_1 = \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}$$

is a basis for $\mathrm{Null}(A)$ and $\dim(\mathrm{Null}(A)) = 2$.

Notice that the first, third, and fourth columns in the RREF of A have leading ones. This means that the first, third, and fourth columns of the matrix A form a basis for the column space of A:

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 1 \end{bmatrix} \right\}$$

is a basis for $\mathrm{Col}(A)$ and $\dim(\mathrm{Col}(A)) = 3$.

The basis for the row space of A is the set containing the transpose of non-zero rows of the RREF of A:

$$B_3 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

is a basis for $\mathrm{Row}(A)$ and $\dim(\mathrm{Row}(A)) = 3$.

Remark
Note that

$$\dim(\mathrm{Col}(A)) = \dim(\mathrm{Row}(A)) = \mathrm{Rank}(A)$$

and that

$$\dim(\mathrm{Null}(A)) = n - \mathrm{Rank}(A)$$

A question appears in Mobius

A question appears in Mobius

Unit 6

Linear Transformations on $\mathbb{R}^n$

6.1 - Linear Mapping

Linear Mappings

Definition of Matrix Transformations

Let $A \in M_{m\times n}(\mathbb{R})$ and $\vec{x} \in \mathbb{R}^n$. Then, $A\vec{x} \in \mathbb{R}^m$. This motivates the following definition.

Definition
For $A \in M_{m\times n}(\mathbb{R})$, we can define a function $L_A : \mathbb{R}^n \to \mathbb{R}^m$ by $L(\vec{x}) = A\vec{x}$, for every $\vec{x} \in \mathbb{R}^n$.
We call L_A the **matrix transformation** corresponding to A.
The domain of L_A is $\mathbb{R}^n$ and the codomain of L_A is $\mathbb{R}^m$.

Example 1

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \end{bmatrix}$.

1. Define the matrix transformation $L_A : \mathbb{R}^3 \to \mathbb{R}^2$.
2. Apply L_A to $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$.

Solution - Part A

We can define $L_A : \mathbb{R}^3 \to \mathbb{R}^2$ by $L(\vec{x}) = A\vec{x}$, for every $\vec{x} \in \mathbb{R}^3$.

So we have:

$$L_A\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ x_1 - x_2 + x_3 \end{bmatrix}$$

Solution - Part B

Applying L_A to $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ gives

$$L_A\left(\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 15 \\ 4 \end{bmatrix}$$

The matrix transformation L_A satisfies the following properties:

Proposition 1: Properties of Matrix Transformations
For $L(\vec{x}) = A\vec{x}$, all $\vec{x}, \vec{y} \in \mathbb{R}^n$, and $k \in \mathbb{R}$:

1. $L_A(\vec{x} + \vec{y}) = L_A(\vec{x}) + L_A(\vec{y})$
2. $L_A(k\vec{x}) = kL_A(\vec{x})$

Proof

For $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $k \in \mathbb{R}$,

1. $L_A(\vec{x} + \vec{y}) = A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} = L_A(\vec{x}) + L_A(\vec{y})$
2. $L_A(k\vec{x}) = A(k\vec{x}) = Ak(\vec{x}) = kA\vec{x} = kL_A(\vec{x})$

$\square$

The proposition tells us that matrix transformations preserve sums and scalar multiplication. This "preservation" property is extremely important in linear algebra - so important that functions with this property are given a special name: linear mappings.

Definition of Linear Mappings

Definition
A **linear mapping** is a type of function that maps vector spaces to vector spaces.
A function $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear mapping if, for $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $k \in \mathbb{R}$,
L1: $L(\vec{x} + \vec{y}) = L(\vec{x}) + L(\vec{y})$, and
L2: $L(k\vec{x}) = kL(\vec{x})$

It follows from this definition that, for $\vec{x}_1, \ldots, \vec{x}_\ell \in \mathbb{R}^n$ and $c_1, \ldots, c_\ell \in \mathbb{R}$,

- $L(c_1\vec{x}_1 + \cdots + c_\ell\vec{x}_\ell) = c_1L(\vec{x}_1) + \cdots + c_\ell L(\vec{x}_\ell)$
- $L(\vec{0}) = \vec{0}$ (taking $k = 0$ in L2)
- $L(-\vec{x}) = -L(\vec{x})$ (taking $k = -1$ in L2)

Thus, linear mappings $L : \mathbb{R}^n \to \mathbb{R}^m$ preserve linear combinations, send zero vectors to zero vectors, and preserve additive inverses.

Comparing the definitions of linear mappings and matrix transformations, we see that matrix transformations are a type of linear mapping: every matrix transformation is a linear mapping.

Example 2

Show that $L : \mathbb{R}^2 \to \mathbb{R}^2$, defined by $L(\vec{x}) = L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ 2x_1 + x_2 \end{bmatrix}$, is a linear mapping.

Solution

To show that L is a linear mapping, we need to show that L satisfies properties L1 and L2.

For property L1: Let $\vec{x}, \vec{y} \in \mathbb{R}^2$. Then,

$$\begin{aligned} L(\vec{x} + \vec{y}) &= L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) \\ &= L\left(\begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} (x_1 + y_1) - (x_2 + y_2) \\ 2(x_1 + y_1) + (x_2 + y_2) \end{bmatrix} \\ &= \begin{bmatrix} x_1 - x_2 \\ 2x_1 + x_2 \end{bmatrix} + \begin{bmatrix} y_1 - y_2 \\ 2y_1 + y_2 \end{bmatrix} \\ &= L(\vec{x}) + L(\vec{y}) \end{aligned}$$

For property L2: Let $\vec{x} \in \mathbb{R}^2$ and $k \in \mathbb{R}$. Then,

$$\begin{aligned} L(k\vec{x}) &= L\left(k\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) \\ &= L\left(\begin{bmatrix} kx_1 \\ kx_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} kx_1 - kx_2 \\ 2kx_1 + kx_2 \end{bmatrix} \\ &= k\begin{bmatrix} x_1 - x_2 \\ 2x_1 + x_2 \end{bmatrix} \\ &= kL(\vec{x}) \end{aligned}$$

Since L satisfies properties L1 and L2, L is a linear mapping.

Note that we can also write L as a matrix transformation:

$$L(\vec{x}) = L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ 2x_1 + x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Remark

It is quite common that the notation $L(x_1, \ldots, x_n)$ is used instead of the notation $L\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)$. In this course, we will be using both notations interchangeably.

Example 3

Show that $L : \mathbb{R}^3 \to \mathbb{R}^2$, defined by $L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_3^2 + 3 \end{bmatrix}$, is **not** a linear mapping.

Solution

To show that L is not a linear mapping, we need to show that either property L1 or property L2 fails.

We will show that property L1 fails, using the vectors $\begin{bmatrix}1\\0\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\\0\end{bmatrix}$.

Since

$$\begin{aligned} L\left(\begin{bmatrix}1\\0\\0\end{bmatrix}+\begin{bmatrix}0\\1\\0\end{bmatrix}\right) &= L\left(\begin{bmatrix}1\\1\\0\end{bmatrix}\right) \\ &= \begin{bmatrix}1+1+0\\0^2+3\end{bmatrix} \\ &= \begin{bmatrix}2\\3\end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} L\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right)+L\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) &= \begin{bmatrix}1+0+0\\0^2+3\end{bmatrix}+\begin{bmatrix}0+1+0\\0^2+3\end{bmatrix} \\ &= \begin{bmatrix}1\\3\end{bmatrix}+\begin{bmatrix}1\\3\end{bmatrix} \\ &= \begin{bmatrix}2\\6\end{bmatrix} \end{aligned}$$

we see that

$$L\left(\begin{bmatrix}1\\0\\0\end{bmatrix}+\begin{bmatrix}0\\1\\0\end{bmatrix}\right) \neq L\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right)+L\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right)$$

so L does not preserve vector additon and therefore is not linear.

> Remark
> When showing that a mapping L is not linear, it is important to show that one of the properties L1 or L2 fails by giving a specific counter-example.

Exercise 1

Show that the mapping L from the above example is not linear, by showing that it does not preserve scalar multiplication (that is, that L2 fails).

A question appears in Mobius

Exercise 2

A question appears in Mobius

Making Connections

Drag the tips of the vectors $\vec{v}_1$ and $\vec{v}_2$ to see how the shape of the parallelogram changes under the linear mapping

$$L(\vec{x}) = L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ 2x_1 + x_2 \end{bmatrix}$$

External resource: https://www.geogebra.org/material/iframe/id/qpfwc2kn/

In the above exercise, you may have noticed how the linear mapping changes the shape (and therefore the area) of the parallelogram induced by the vectors $\vec{v}_1$ and $\vec{v}_2$. Although the example is in $\mathbb{R}^2$, the behaviour of linear mappings in $\mathbb{R}^n$ is analogous. We often think of linear mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$ as transformations of the space. In later lessons, we will see that such mappings scale areas in a predictable way.

6.2 - Matrix of a Linear Mapping

Matrix of a Linear Mapping

Motivating Example

Last lesson, we learned about matrix transformations and linear mappings and saw that every matrix transformation is a linear mapping.

In this lesson, we will see that the converse is also true: every linear mapping is also a matrix transformation.

Let's see an example to get us started.

Example 1

Let $L\colon \mathbb{R}^3 \to \mathbb{R}^2$ be a linear mapping such that

$$L\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}1\\-1\end{bmatrix}, \quad L\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\3\end{bmatrix}, \quad L\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\1\end{bmatrix}$$

Find $L\left(\begin{bmatrix}2\\-3\\1\end{bmatrix}\right)$.

Solution

$$\begin{aligned}
L\left(\begin{bmatrix}2\\-3\\1\end{bmatrix}\right) &= L\left(2\begin{bmatrix}1\\0\\0\end{bmatrix} + (-3)\begin{bmatrix}0\\1\\0\end{bmatrix} + 1\begin{bmatrix}0\\0\\1\end{bmatrix}\right)\\
&= 2L\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) + (-3)L\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) + 1L\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right)\\
&= 2\begin{bmatrix}1\\-1\end{bmatrix} + (-3)\begin{bmatrix}1\\3\end{bmatrix} + 1\begin{bmatrix}0\\1\end{bmatrix}\\
&= \begin{bmatrix}-1\\-10\end{bmatrix}
\end{aligned}$$

Exercise 1

A question appears in Mobius

Example 2

Let $L : \mathbb{R}^2 \to \mathbb{R}^4$ be a linear mapping such that $L\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}1\\2\\3\\4\end{bmatrix}$ and $L\left(\begin{bmatrix}2\\3\end{bmatrix}\right) = \begin{bmatrix}1\\4\\0\\-1\end{bmatrix}$.

Find $L\left(\begin{bmatrix}3\\5\end{bmatrix}\right)$.

Solution

$$\begin{aligned} L\left(\begin{bmatrix}3\\5\end{bmatrix}\right) &= L\left(\begin{bmatrix}1\\2\end{bmatrix} + \begin{bmatrix}2\\3\end{bmatrix}\right) \\ &= L\left(\begin{bmatrix}1\\2\end{bmatrix}\right) + L\left(\begin{bmatrix}2\\3\end{bmatrix}\right) \\ &= \begin{bmatrix}1\\2\\3\\4\end{bmatrix} + \begin{bmatrix}1\\4\\0\\-1\end{bmatrix} \\ &= \begin{bmatrix}2\\6\\3\\3\end{bmatrix} \end{aligned}$$

As this example shows, knowing how L transforms the vectors $\begin{bmatrix}1\\2\end{bmatrix}$ and $\begin{bmatrix}2\\3\end{bmatrix}$ is enough to tell us how L transforms any linear combination of $\begin{bmatrix}1\\2\end{bmatrix}$ and $\begin{bmatrix}2\\3\end{bmatrix}$.

Exercise 2

A question appears in Mobius

In general, for a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$, knowing $L(\vec{x}_1), \dots, L(\vec{x}_k)$ for $\vec{x}_1, \dots, \vec{x}_k \in \mathbb{R}^n$ means that we can compute $L(\vec{x})$ for any $\vec{x} \in \operatorname{Span}\{\vec{x}_1, \dots, \vec{x}_k\}$.

In particular, if $\{\vec{w}_1, \dots, \vec{w}_n\}$ is a basis for $\mathbb{R}^n$ and we know $L(\vec{w}_1), \dots, L(\vec{w}_n)$, then we can compute $L(\vec{w})$ for any $\vec{w} \in \mathbb{R}^n$.

Indeed, in our last example, $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$, thus for $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$,

$$\left[\begin{array}{cc|c} 1 & 2 & x_1 \\ 2 & 3 & x_2 \end{array}\right] \underset{R_2-2R_1}{\sim} \left[\begin{array}{cc|c} 1 & 2 & x_1 \\ 0 & -1 & x_2 - 2x_1 \end{array}\right] \underset{R_1+2R_2}{\sim} \left[\begin{array}{cc|c} 1 & 0 & -3x_1 + 2x_2 \\ 0 & -1 & x_2 - 2x_1 \end{array}\right] \underset{-R_2}{\sim} \left[\begin{array}{cc|c} 1 & 0 & -3x_1 + 2x_2 \\ 0 & 1 & 2x_1 - x_2 \end{array}\right]$$

so

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (-3x_1 + 2x_2) \begin{bmatrix} 1 \\ 2 \end{bmatrix} + (2x_1 - x_2) \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

and

$$\begin{aligned} L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) &= L\left((-3x_1 + 2x_2) \begin{bmatrix} 1 \\ 2 \end{bmatrix} + (2x_1 - x_2) \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right) \\ &= (-3x_1 + 2x_2) L\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix} \right) + (2x_1 - x_2) L\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix} \right) \\ &= (-3x_1 + 2x_2) \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + (2x_1 - x_2) \begin{bmatrix} 1 \\ 4 \\ 0 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} -x_1 + x_2 \\ 2x_1 \\ -9x_1 + 6x_2 \\ -14x_1 + 9x_2 \end{bmatrix} \end{aligned}$$

We can rewrite this as the following matrix transformation:

$$L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} -x_1 + x_2 \\ 2x_1 \\ -9x_1 + 6x_2 \\ -14x_1 + 9x_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & 0 \\ -9 & 6 \\ -14 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Starting from a basis of $\mathbb{R}^2$, we were able to rewrite the linear mapping L as a matrix transformation. As we will see in the next section, we will always be able to do this.

Remark

In the example above, we chose the basis $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\3\end{bmatrix}\right\}$ of $\mathbb{R}^2$, but we could have chosen **any** basis of $\mathbb{R}^2$ so to understand the action of L.

Standard Matrix

Theorem 1

If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear mapping, then L is a matrix transformation with corresponding matrix

$$[L] = [L(\vec{e}_1) \cdots L(\vec{e}_n)] \in M_{m\times n}(\mathbb{R})$$

which is called the **standard matrix** of L.

Proof

Let $\vec{x} = [x_1 \cdots x_n]^T \in \mathbb{R}^n$. Then

$$\begin{aligned}
L(\vec{x}) &= L(x_1\vec{e}_1 + \cdots + x_n\vec{e}_n) \\
&= x_1L(\vec{e}_1) + \cdots + x_nL(\vec{e}_n) && \text{since } L \text{ is linear} \\
&= [L(\vec{e}_1) \cdots L(\vec{e}_n)] \begin{bmatrix}x_1\\ \vdots \\ x_n\end{bmatrix} \\
&= [L]\vec{x}
\end{aligned}$$

□

It follows from the definition of a matrix transformation that we introduced in the lesson on linear mappings that $L : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation if and only if it is a linear mapping.

Example 3

Let $L : \mathbb{R}^2 \to \mathbb{R}^4$ be a linear mapping such that $L\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}1\\2\\3\\4\end{bmatrix}$ and $L\left(\begin{bmatrix}2\\3\end{bmatrix}\right) = \begin{bmatrix}1\\4\\0\\-1\end{bmatrix}$.

Find the standard matrix of L.

Solution

The calculations following Exercise 2 led us to the conclusion that

$$L\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = \begin{bmatrix}-x_1 + x_2\\ 2x_1 \\ -9x_1 + 6x_2 \\ -14x_1 + 9x_2\end{bmatrix}$$

To find the standard matrix of L, we calculate $L(\vec{e}_1)$ and $L(\vec{e}_2)$:

$$L\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}-1\\2\\-9\\-14\end{bmatrix},\quad L\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}1\\0\\6\\9\end{bmatrix}$$

Therefore, the standard matrix of L is

$$\begin{bmatrix}-1&1\\2&0\\-9&6\\-14&9\end{bmatrix}$$

Example 4

Let $L\colon \mathbb{R}^2\to\mathbb{R}^4$ be a linear mapping such that $L\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}-1\\2\\-9\\-14\end{bmatrix}$ and $L\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}1\\0\\6\\9\end{bmatrix}$.

Find a matrix A such that $L(\vec{x})=A\vec{x}$ for any $\vec{x}\in\mathbb{R}^2$.

Solution

Notice that

$$\begin{aligned}L\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)&=x_1L\left(\begin{bmatrix}1\\0\end{bmatrix}\right)+x_2L\left(\begin{bmatrix}0\\1\end{bmatrix}\right)\\&=x_1\begin{bmatrix}-1\\2\\-9\\-14\end{bmatrix}+x_2\begin{bmatrix}1\\0\\6\\9\end{bmatrix}\\&=\begin{bmatrix}-1&1\\2&0\\-9&6\\-14&9\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}\end{aligned}$$

Therefore,

$$A=\begin{bmatrix}-1&1\\2&0\\-9&6\\-14&9\end{bmatrix}$$

Remark
Notice that **any** $m\times n$ matrix A is the standard matrix of a linear mapping. This is not a coincidence. As we stated at the start of this lesson, every matrix transformation is a linear mapping and every linear mapping is also a matrix transformation.

Exercise 3

A question appears in Mobius

Making Connections

In this exercise, you will construct your own linear mapping and see its standard matrix.

Drag the tips of the vectors $L(\vec{e}_1)$ and $L(\vec{e}_2)$ on the plane to define your new linear mapping and see how its standard matrix $[L]$ changes as you change $L(\vec{e}_1)$ and $L(\vec{e}_2)$.

External resource: https://www.geogebra.org/material/iframe/id/g9sxftwy/

Remark
The interactive example shows that it is enough to know how the mapping acts on the standard basis vectors. This is a particular case of the following: if $L : \mathbb{R}^n \to \mathbb{R}^m$, $\{\vec{w}_1, \ldots, \vec{w}_n\}$ is a basis for $\mathbb{R}^n$, and we know $L(\vec{w}_1), \ldots, L(\vec{w}_n)$, then we can compute $L(\vec{w})$ for any $\vec{w} \in \mathbb{R}^n$.

Exercise 4

A question appears in Mobius

6.3 - Composition of Transformations and Matrix Products

Operations on Linear Mappings

We will now study linear mappings more algebraically and see that linear mappings follow the same rules as vectors in $\mathbb{R}^n$ and $m \times n$ matrices.

Addition and Scalar Multiplication of Linear Mappings

We start by defining the addition and scalar multiplication of linear mapping:

Definitions
Let $S, T : \mathbb{R}^n \to \mathbb{R}^m$ be linear mappings and let $k \in \mathbb{R}$. The **sum** of S and T, denoted by $S + T$, is defined by

$$(S + T)(\vec{x}) = S(\vec{x}) + T(\vec{x})$$

for all $\vec{x} \in \mathbb{R}^n$.
The **scalar multiple** of T, denoted kT, is defined by

$$(kT)(\vec{x}) = kT(\vec{x})$$

for all $\vec{x} \in \mathbb{R}^n$.

We note the following useful result.

Theorem 1
If $S, T : \mathbb{R}^n \to \mathbb{R}^m$ are linear mappings, then so are $(S + T)$ and kT for any $k \in \mathbb{R}$.

Proof

Let $S, T : \mathbb{R}^n \to \mathbb{R}^m$ be linear mappings and let $k \in \mathbb{R}$.

First, we will show that $(S + T)$ is linear by proving that it satifies properties L1 and L2.

For property L1: Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. We have

$$\begin{aligned}
(S + T)(\vec{x} + \vec{y}) &= S(\vec{x} + \vec{y}) + T(\vec{x} + \vec{y}) && \text{by definition of } S + T \\
&= S(\vec{x}) + S(\vec{y}) + T(\vec{x}) + T(\vec{y}) && \text{since } S \text{ and } T \text{ are linear} \\
&= (S + T)(\vec{x}) + (S + T)(\vec{y}).
\end{aligned}$$

For property L2: Let $\ell \in \mathbb{R}$. We have

$$\begin{aligned}
(S + T)(\ell\vec{x}) &= S(\ell\vec{x}) + T(\ell\vec{x}) && \text{by definition of } S + T \\
&= \ell S(\vec{x}) + \ell T(\vec{x}) && \text{since } S \text{ and } T \text{ are linear} \\
&= \ell(S(\vec{x}) + T(\vec{x})) \\
&= \ell(S + T)(\vec{x})
\end{aligned}$$

Since $(S + T)$ satisfies properties L1 and L2, $(S + T)$ is a linear mapping.

Next, we will show that kT is linear by proving that it satifies L1 and L2.

For property L1: Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. We have

$$\begin{aligned}(kT)(\vec{x}+\vec{y}) &= kT(\vec{x}+\vec{y}) && \text{by definition of } kT\\ &= k(T(\vec{x})+T(\vec{y})) && \text{since } T \text{ is linear}\\ &= kT(\vec{x})+kT(\vec{y})\\ &= (kT)(\vec{x})+(kT)(\vec{y})\end{aligned}$$

For property L2: Let $\ell \in \mathbb{R}$. We have

$$\begin{aligned}(kT)(\ell\vec{x}) &= kT(\ell\vec{x})\\ &= k(\ell T(\vec{x})) \quad \text{since T is linear}\\ &= \ell(k(T(\vec{x}))\\ &= \ell(kT)(\vec{x})\end{aligned}$$

Since kT satisfies properties L1 and L2, kT is a linear mapping.

□

Example 1

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ x_1 + x_2 \end{bmatrix}$, and

let $S : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by $S\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ x_1 \\ x_2 \end{bmatrix}$.

Determine $S + T$ and $2T$.

Solution

Using the theorem, we have

$$\begin{aligned}(S+T)\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= S\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)\\ &= \begin{bmatrix} x_1 - x_2 \\ x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} x_1 \\ x_2 \\ x_1 + x_2 \end{bmatrix}\\ &= \begin{bmatrix} x_1 - x_2 + x_1 \\ x_1 + x_2 \\ x_2 + x_1 + x_2 \end{bmatrix}\\ &= \begin{bmatrix} 2x_1 - x_2 \\ x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix}\end{aligned}$$

We also have

$$\begin{aligned}2T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= 2\begin{bmatrix} x_1 \\ x_2 \\ x_1 + x_2 \end{bmatrix}\\ &= \begin{bmatrix} 2x_1 \\ 2x_2 \\ 2x_1 + 2x_2 \end{bmatrix}\end{aligned}$$

Making Connections

The mappings, S, T, $S+T$, and $2T$, as defined in the previous example, map vectors in $\mathbb{R}^2$ to $\mathbb{R}^3$. But how does this look, geometrically?

Select the tip of the vector $\vec{u}$ and drag it to any position to observe how it is transformed by S, T and $S+T$, and $2T$.

External resource: https://www.geogebra.org/material/iframe/id/pdhb9anc/

Exercise 1

A question appears in Mobius

Composition of Linear Mappings

In addition to adding and scalar multiplying linear mappings, we can define their composition as follows:

Definition
Let $T:\mathbb{R}^n \to \mathbb{R}^m$ and $S:\mathbb{R}^m \to \mathbb{R}^p$ be linear mappings. The **composition** of S and T, $S\circ T:\mathbb{R}^n \to \mathbb{R}^p$ is defined by

$$(S\circ T)(\vec{x}) = S(T(\vec{x}))$$

for all $\vec{x}\in\mathbb{R}^n$.

Note that, since $T:\mathbb{R}^n \to \mathbb{R}^m$ and $S:\mathbb{R}^m \to \mathbb{R}^p$, the composition sends:

- $\vec{x}\in\mathbb{R}^n$ to $T(\vec{x})\in\mathbb{R}^m$ and
- $T(\vec{x})\in\mathbb{R}^m$ to $S(T(\vec{x}))\in\mathbb{R}^p$,

so indeed, $S\circ T:\mathbb{R}^n \to \mathbb{R}^p$.

We note the following useful result:

Theorem 2: Composition of Linear Mappings
If $T : \mathbb{R}^n \to \mathbb{R}^m$and $S : \mathbb{R}^m \to \mathbb{R}^p$ are linear mappings, then $(S \circ T)$ is also a linear mapping.

Proof

Let $T : \mathbb{R}^n \to \mathbb{R}^m$and $S : \mathbb{R}^m \to \mathbb{R}^p$ be linear mappings.

We will show that $(S \circ T)$ is linear by proving that it satifies L1 and L2.

For property L1: Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. We have

$$\begin{aligned}
(S \circ T)(\vec{x} + \vec{y}) &= S(T(\vec{x} + \vec{y})) && \text{by definition of } S \circ T \\
&= S(T(\vec{x}) + T(\vec{y})) && \text{since } T \text{ is linear} \\
&= (S(T(\vec{x})) + (S(T(\vec{y})) && \text{since } S \text{ is linear} \\
&= (S \circ T)(\vec{x}) + (S \circ T)(\vec{y})
\end{aligned}$$

For property L2: Let $k \in \mathbb{R}$. We have

$$\begin{aligned}
(S \circ T)(k\vec{x}) &= S((T(k\vec{x})) && \text{by definition of } S \circ T \\
&= S(k(T(\vec{x})) && \text{since T is linear} \\
&= k(S(T(\vec{x})) && \text{since S is linear} \\
&= k(S \circ T)(\vec{x})
\end{aligned}$$

Since $(S \circ T)$ satisfies properties L1 and L2, $(S \circ T)$ is a linear mapping.

□

Example 2

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ and $S : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 \\ 3x_2 \\ x_1 + x_2 \end{bmatrix}$ and $S\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_3 \\ x_1 \\ x_2 \end{bmatrix}$.

Determine $(S \circ T) : \mathbb{R}^2 \to \mathbb{R}^3$.

Solution

Then $(S \circ T) : \mathbb{R}^2 \to \mathbb{R}^3$ is given by

$$\begin{aligned}
(S \circ T)\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= S\left(T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)\right) \\
&= S\left(\begin{bmatrix} 2x_1 \\ 3x_2 \\ x_1 + x_2 \end{bmatrix}\right) \\
&= \begin{bmatrix} x_1 + x_2 \\ 2x_1 \\ 3x_2 \end{bmatrix}
\end{aligned}$$

Making Connections

The mapping $S \circ T$, as defined in the previous example, map vectors in R^2 to R^3. How does this look, geometrically?

Select the tip of the vector $\vec{u}$ and drag it to any position to observe how it is transformed by $S \circ T$.

External resource: https://www.geogebra.org/material/iframe/id/wj7j2xpp/

Exercise 2

A question appears in Mobius

Operations on Linear Mappings and Standard Matrices of Linear Mappings

As you might expect, the operations on linear mappings that we saw in the previous sections carry over to the matrices of the linear mappings. The next theorem shows us how to construct the standard matrices of these new mappings.

Therorem 3: Properties of Standard Matrices
Let $T, R : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$ be linear mappings and $k \in \mathbb{R}$. Then,

1. $[T + R] = [T] + [R]$
2. $[kT] = k[T]$
3. $[S \circ T] = [S][T]$

We will prove property 3 of the theorem.

Proof of Property 3

Let $\vec{x} \in \mathbb{R}^n$. Then,

$$\begin{aligned}
[S \circ T]\vec{x} &= (S \circ T)(\vec{x}) \\
&= S(T(\vec{x})) \\
&= S([T]\vec{x}) \\
&= [S]([T]\vec{x}) \\
&= [S][T]\vec{x}
\end{aligned}$$

Thus, $[S \circ T]\vec{x} = [S][T]\vec{x}$ for any $\vec{x} \in \mathbb{R}^n$ so we have $[S \circ T] = [S][T]$ by the **Matrices Equal Theorem** stated in the lesson on matrix-vector multiplication.

We see that if $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$ are linear mappings, then the standard matrix for $S \circ T$ is the product of the standard matrix of S and the standard matrix of T, in that order. Thus matrix multiplication can be viewed as composing linear mappings.

Recall that for linear mappings $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$, $[T] \in M_{m\times n}(\mathbb{R})$ and $[S] \in M_{p\times m}(\mathbb{R})$. Since $S \circ T : \mathbb{R}^n \to \mathbb{R}^p$, $[S \circ T] \in M_{p\times n}(\mathbb{R})$. Indeed,

$$\begin{array}{ccccc} [S \circ T] & = & [S] & & [T] \\ \uparrow & & \uparrow & & \uparrow \\ p \times n & & p \times m & & m \times n \end{array}$$

so the matrix product is defined.

□

Example 3

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ and $S : \mathbb{R}^2 \to \mathbb{R}^2$ be linear mappings whose standard matrices are

$$[T] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad [S] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Calculate $[S \circ T]$ and $[T \circ S]$.

Solution

We have

$$[S \circ T] = [S][T] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$
$$[T \circ S] = [T][S] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

This example also illustrates the fact that $[S \circ T] \neq [T \circ S]$ and hence $S \circ T \neq T \circ S$. This fits with our knowledge that the composition of functions is not commutative.

Exercise 3

A question appears in Mobius

Invertible Linear Mappings

The Identity Mapping Viewed as the Composition of Two Linear Mappings

Here, we will study a particular type of linear mapping which has the property of being invertible. We motivate this definition by an example.

Example 1

Consider the linear mappings $T, S : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + x_2 \\ x_1 + x_2 \end{bmatrix}$$

$$S\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ -x_1 + 2x_2 \end{bmatrix}$$

Show that the standard matrices of T and S satisfy $[S]^{-1} = [T]$ and $[T]^{-1} = [S]$.

Solution

The standard matrices of T and S are

$$[T] = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

and

$$[S] = \begin{bmatrix} S(\vec{e}_1) & S(\vec{e}_2) \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

respectively. Therefore,

$$[S \circ T] = [S][T] = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$[T \circ S] = [T][S] = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

From this example, we see that $[S]^{-1} = [T]$ and that $[T]^{-1} = [S]$.

Making Connections

The mappings, S and T, as defined in the previous example, map vectors in $\mathbb{R}^2$ to $\mathbb{R}^2$. How does this look, geometrically? Select the tip of the vector $\vec{u}$ and drag it to any position to observe how it is transformed.

External resource: https://www.geogebra.org/material/iframe/id/fsz8mwrh/

Before introducing our main definition, we define the identity mapping.

Definition
The linear mapping $\mathrm{Id} : \mathbb{R}^n \to \mathbb{R}^m$ defined by $\mathrm{Id}(\vec{x}) = \vec{x}$ for every $\vec{x} \in \mathbb{R}^n$ is called the **identity mapping**. The identity matrix is such that $[\mathrm{Id}] = I$, where I is the $n \times n$ identity matrix.

Now, our main definition.

Definition
If $S, T : \mathbb{R}^n \to \mathbb{R}^n$ are linear mappings such that

$$S \circ T = \mathrm{Id} = T \circ S$$

then T is **invertible** and $T^{-1} = S$; S is also invertible with $S^{-1} = T$.

Geometrically, if $T^{-1} = S$, then we can think of S as "undoing" what T does.

Note that $T(T^{-1}(\vec{x})) = \vec{x}$ for any $\vec{x} \in \mathbb{R}^n$, therefore for $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have

$$\begin{aligned} \vec{x} + \vec{y} &= T(T^{-1}(\vec{x})) + T(T^{-1}(\vec{y})) \\ &= T(T^{-1}(\vec{x}) + T^{-1}(\vec{y})) && \text{since } T \text{ is linear} \end{aligned}$$

We will use this observation to prove the following result, which states that the inverse of a linear mapping is also a linear mapping.

Theorem 4: Invertible Linear Mapping
Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear mapping. Then $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is also a linear mapping.

Proof

We will show that T^{-1} is linear by showing that it satisfies properties L1 and L2.
For property L1: Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. Then,

$$\begin{aligned} T^{-1}(\vec{x} + \vec{y}) &= T^{-1}(T(T^{-1}(\vec{x}) + T^{-1}(\vec{y}))) && \text{by the observation} \\ &= T^{-1}(\vec{x}) + T^{-1}(\vec{y}) \end{aligned}$$

For property L2: Let $\vec{x} \in \mathbb{R}^n$ and $k \in \mathbb{R}$. Then,

$$\begin{aligned} \vec{x} &= T(T^{-1}(\vec{x})) \\ k\vec{x} &= kT(T^{-1}(\vec{x}) \\ &= T(kT^{-1}(\vec{x})) \qquad (*) \end{aligned}$$

which means that

$$\begin{aligned} T^{-1}(k\vec{x}) &= T^{-1}(T(kT^{-1}(\vec{x}))) \quad \text{by } (*) \\ &= kT^{-1}(\vec{x}) \end{aligned}$$

Since T^{-1} satisfies properties L1 and L2, it is a linear mapping.

□

The following result relates inverse mappings to matrix inverses.

Theorem 5
If $S, T : \mathbb{R}^n \to \mathbb{R}^n$ are linear mappings, then S is the inverse of T if and only if $[S]$ is the inverse of $[T]$.

Proof

We have

$$\begin{aligned} \text{S is the inverse of T} &\Leftrightarrow S \circ T = \text{Id} = T \circ S \\ &\Leftrightarrow [S \circ T] = [\text{Id}] = [T \circ S] \\ &\Leftrightarrow [S][T] = I = [T][S] \\ &\Leftrightarrow [S] \text{ is the inverse of } [T] \end{aligned}$$

From the theorem, it follows that for an invertible linear mapping $T : \mathbb{R}^n \to \mathbb{R}^n$,

$$[T^{-1}] = [T]^{-1}$$

□

Example 2

Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a linear mapping defined by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \end{bmatrix}$$

Determine L^{-1}.

Solution

We find that the standard matrix of L is

$$[L] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Since $[L^{-1}] = [L]^{-1}$, we find $[L]^{-1}$ by row reducing the system:

$$\left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & 1/2 & 1/2 \\ 0 & 1 & -1/2 & 1/2 \end{array}\right]$$

We find that

$$[L^{-1}] = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

and therefore the mapping $L^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$L^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{2}x_1 + \frac{1}{2}x_2 \\ -\frac{1}{2}x_1 + \frac{1}{2}x_2 \end{bmatrix}$$

A question appears in Mobius

A question appears in Mobius

6.4 - Function Basics

Function Basics

Introduction

In this section, we will review some basic facts and properties of functions that will be useful to us when we study linear transformations. Let's start with the formal definition of a function.

> Definitions
> A **function** $f : A \to B$ is a relation between sets A and B that associates to each element of A exactly one element of B. The set A is called the **domain** of f and the set B is called the **codomain** of f.

We commonly view functions as maps from a set A to a set B. You are probably most familiar with functions $f : \mathbb{R} \to \mathbb{R}$ which map real numbers to real numbers. Later, we will see that linear transformations are a type of function that map vector spaces to vector spaces.

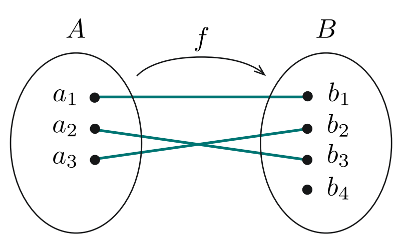
In the next sections, we will discuss some characteristics that a function might have.

Injective Functions

Let's begin by looking at the domain of a linear transformation $f : A \to B$. We are interested in where f sends each vector in the domain.

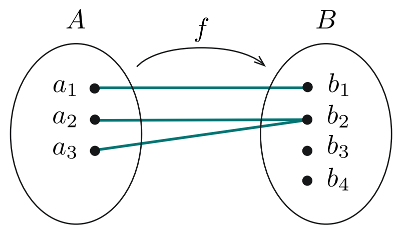
> Definition
> A function $f : A \to B$ is **injective** (or one-to-one) if $f(a_1) = f(a_2)$ implies that $a_1 = a_2$ for all $a_1, a_2 \in A$.

An **injective** function is one that maps each element of the domain to a unique element of the codomain. Let's see a visual representation of this.



As the picture shows, each element of the domain is sent to a unique element in the codomain - no two elements in the domain "share" an element of the codomain.

The next illustration shows a function which is **not injective**:



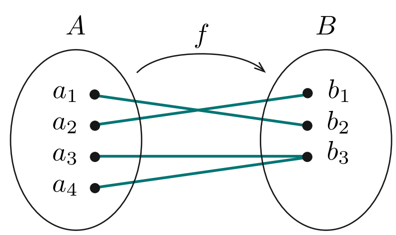
This function is not injective since the elements a_2 and a_3 are both mapped to the element b_2 in the codomain, so $f(a_2) = f(a_3)$ but $a_2 \neq a_3$.

Surjective Functions

Now, let's focus on the codomain of a function $f : A \to B$. We are interested in whether every element of the codomain can be reached by f.

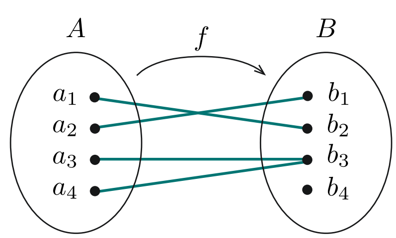
> Definition
> Let $f : A \to B$ be a function. f is **surjective** (or onto) if for every $b \in B$ there exists an $a \in A$ such that $f(a) = b$.

A **surjective** function is one that maps "onto" every vector in the codomain at least once. Let's see a visual representation of this.



As the picture shows, each element of the codomain is "reached" by an element in the domain - no element in the codomain is left untouched by the function.

The next illustration shows a function which is **not surjective**:



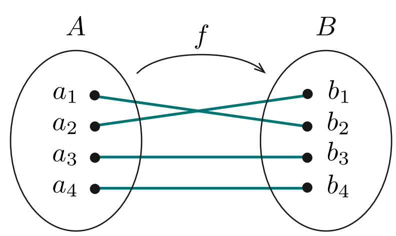
This function is not surjective since the element b_4 has not been mapped to by the function.

Bijective Functions

When a function is both injective and surjective, we give it a special name.

Definition
A function $f : A \to B$ is **bijective** (or invertible) if it is both injective and surjective.

A **bijective** function is one which maps each element in the domain to a unique element in the codomain and which maps on to each element of the codomain exactly once. Let's see a visual representation of this.



As the image shows, every element of the codomain is mapped to exactly once and every element of the domain maps on to a unique element in the codomain.

Remark
Note that although the concept of a bijective function relates the concepts of injective and surjective functions, it is possible for a function to be:

- injective but not surjective, or
- surjective but not injective.

Try to come up with examples of such functions!

Exercise

A question appears in Mobius

6.5 - Special Spaces (Kernel, Range)

Special Spaces of a Linear Mapping

Kernel of a Linear Mapping

Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. The **kernel** of L is the set of vectors in $\mathbb{R}^n$ which are mapped to the zero vector in $\mathbb{R}^m$ by L. Formally, we have the following definition.

Definiton
The **kernel** of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ is

$$\ker(L) = \{\vec{x} \in \mathbb{R}^n \mid L(\vec{x}) = \vec{0}_{\mathbb{R}^m}\} \subseteq \mathbb{R}^n$$

Example 1

Let $L : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear mapping defined by $L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix}$.

Determine whether the vector $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ is in $\ker(L)$.

Solution

The vector $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ is in $\ker(L)$ since

$$L\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1-1 \\ -1+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Exercise 1

A question appears in Mobius

Range of a Linear Mapping

Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. The **range** of L is the set of vectors in $\mathbb{R}^m$ which are reached by L. Formally, we have the following definition.

Definition
The **range** of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ is

$$\text{Range}(L) = \{L(\vec{x}) \mid \vec{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

Exercise 2

A question appears in Mobius

Exercise 3

For a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$, prove that $\ker(L)$ is a subspace of $\mathbb{R}^n$ and that $\text{Range}(L)$ is a subspace of $\mathbb{R}^m$.

A question appears in Mobius

Finding a Basis for the Kernel and Range of a Linear Mapping

As was shown in the exercise, the kernel and range of a linear mapping form subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. This means that we can find a basis for each of these subspaces. The procedure is quite similar to the procedure we learned in the lesson on subspaces. Let's see how it works.

Example 2

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear mapping with standard matrix

$$[T] = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

1. Find a basis for $\ker(T)$.
2. Find a basis for $\text{Range}(T)$.

Solution - Part A

If $\vec{x} \in \ker(T)$, then $[T]\vec{x} = \vec{0}$. Elements of the kernel are therefore solutions to the homogeneous system $[T]\vec{x} = \vec{0}$.

Since

$$\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

we have

$$\vec{x} = s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad s, t \in \mathbb{R}$$

so

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a basis for $\ker(T)$.

Solution - Part B

For $\vec{y} \in \text{range}(T)$, there must be an $\vec{x} \in \mathbb{R}^3$ so $[T]\vec{x} = \vec{y}$ must be consistent, so $\vec{y} \in \text{Col}([T])$. Observing the RREF of $[T]$ from above, we see that there is a leading one in the first column. Therefore,

$$\left\{ \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} \right\}$$

is a basis for $\text{Range}(T)$.

You might have noticed that the kernel and range of a linear mapping L are closely related to the nullspace and column space of $[L]$. The following theorem formalizes this intuition.

Theorem 1
For a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$

1. $\ker(L) = \text{Null}([L])$
2. $\text{Range}(L) = \text{Col}([L])$

The theorem tells us that finding the kernel and range of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ is the same as finding the nullspace and column space of $[L]$, respectively. Fortunately, we have a lot of practice doing this.

Example 3

Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear mapping defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix}$.

1. Find a basis for $\ker(T)$.
2. Find a basis for Range(T).

Solution - Part A

We start by finding $[T]$:

$$[T] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

and since

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

the solution of $T(\vec{x}) = \vec{0}$ is

$$\vec{x} = t \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

and therefore

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

is a basis for $\ker(T)$.

Solution - Part B

Since the REF of T has leading entries in the first two columns,

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

is a basis for Range(T).

A question appears in Mobius

A question appears in Mobius

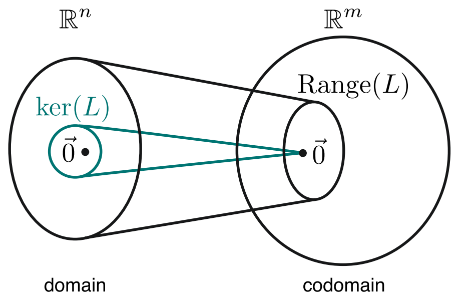
Injective, Surjective, and Bijective Linear Mappings

Introduction

A linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ maps vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$. The set $\mathbb{R}^n$ is called the **domain** of L, and the set $\mathbb{R}^m$ is called the **codomain** of L. As we have seen in previous examples,

- the kernel of L is the set of vectors from the domain which are mapped to the zero vector in $\mathbb{R}^m$, and
- the range of L is the set of vectors in $\mathbb{R}^m$ which are reached by L, and is not always equal to its codomain.

If we were to draw a rough sketch of what's going on, we might come up with something like this:



In the next sections, we will discuss some characteristics that a linear mapping might have.

Injective Linear Mappings

Let's begin by looking at the domain of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$. We are interested in where L sends each vector in the domain.

Definition
Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. L is **injective** (or one-to-one) if $L(\vec{x}_1) = L(\vec{x}_2)$ implies that $\vec{x}_1 = \vec{x}_2$ for all $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$.

An injective linear mapping maps each vector in the domain to a unique vector in the codomain. As it turns out, determining whether a linear mapping is injective is very closely related to the kernel of the mapping.

Theorem 2: Injective Linear Mapping Theorem
Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. L is **injective** if and only if $\ker(L) = \{\vec{0}\}$.

Proof

This is an "if and only if" statement, so we need to show two implications.

First, we show that if L is injective, then $\ker(L) = \{\vec{0}\}$.

Suppose that L is injective. We want to show that the only element in the kernel of L is the zero vector.

Let $\vec{x} \in \ker(L)$. Then by definition of the kernel, $L(\vec{x}) = \vec{0}$. We also have $L(\vec{0}) = \vec{0}$, that is, $L(\vec{x}) = L(\vec{0})$.

Since L is injective, this means that $\vec{x} = \vec{0}$. Hence, any vector in the kernel of L must be equal to the zero vector.

Second, we show that if $\ker(L) = \{\vec{0}\}$, then L is injective.

Suppose that $\ker(L) = \{\vec{0}\}$ and let $\vec{x}, \vec{y} \in \mathbb{R}^n$ be such that $L(\vec{x}) = L(\vec{y})$. We want to show that $\vec{x} = \vec{y}$.

We have

$$\begin{aligned} L(\vec{x}) &= L(\vec{y}) \\ L(\vec{x}) - L(\vec{y}) &= \vec{0} \\ L(\vec{x} - \vec{y}) &= \vec{0} && \text{since } L \text{ is linear} \end{aligned}$$

As $L(\vec{x} - \vec{y}) = \vec{0}$, this means that $\vec{x} - \vec{y} \in \ker(L)$.

Furthermore, since the only element in $\ker(L)$ is $\vec{0}$, we have $\vec{x}-\vec{y}=\vec{0}$ and therefore $\vec{x}=\vec{y}$, which shows that L is injective.

□

Remark
The fact that $\ker(L)=\{\vec{0}\}$ implies that $[L]\vec{x}=\vec{0}$ has only the trivial solution, so any REF of $[L]$ has a leading entry in each column and hence $\operatorname{rank}[L]=n$. In other words, the rank of $[L]$ is equal to the **number of columns** of $[L]$.

Example 1

Determine whether the linear mapping $L\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=\begin{bmatrix}x_1\\x_2-x_1\\x_2\end{bmatrix}$ is injective.

Solution

Since

$$[L]=\begin{bmatrix}1&0\\-1&1\\0&1\end{bmatrix}\sim\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}$$

we see that $\operatorname{rank}([L])=2$, which is equal to the number of columns of $[L]$, so $L(\vec{x})=\vec{0}$ has only the trivial solution and hence $\ker(L)=\{\vec{0}\}$.

Thus, L is injective by the Injective Linear Mapping Theorem.

Example 2

Determine whether the linear mapping $S\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=\begin{bmatrix}x_1+x_2\\2x_1+2x_2\end{bmatrix}$ is injective.

Solution

Since

$$[S]=\begin{bmatrix}1&1\\2&2\end{bmatrix}\sim\begin{bmatrix}1&1\\0&0\end{bmatrix}$$

we have that $\operatorname{rank}([S])=1$, which is less than the number of columns of $[S]$. This means that $S(\vec{x})=\vec{0}$ has non-trivial solutions and hence $\ker(S)\neq\{\vec{0}\}$. Thus, by the theorem above, S is not injective.

Exercise 1

A question appears in Mobius

Surjective Linear Mappings

Now, let's focus on the codomain of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$. We are interested in whether every element of the codomain can be reached by L.

Definition
Let $L : \mathbb{R}^n \to \mathbb{R}^m$be a linear mapping. L is **surjective** (or onto) if for every $\vec{y} \in \mathbb{R}^m$ there exists an $\vec{x} \in \mathbb{R}^n$ such that $L(\vec{x}) = \vec{y}$.

A **surjective** linear mapping is one that maps "onto" every vector in the codomain at least once. As it turns out, determining whether a linear mapping is surjective is very closely related to the range of the mapping.

Theorem 3
Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. L is surjective if and only if $\text{Range}(L) = \mathbb{R}^m$.

Remark
If a linear mapping $L\colon \mathbb{R}^n \to \mathbb{R}^m$ is surjective, then the equation $[L]\vec{x} = \vec{b}$ has a solution for every $\vec{b} \in \mathbb{R}^m$. The properties of the rank tell us that this happens if and only if $\text{rank}[L] = m$. In other words, the rank of $[L]$ is equal to the **number of rows** of $[L]$.

Example 3

Determine whether the linear mapping $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 - x_3 \\ x_2 + x_3 \end{bmatrix}$is surjective.

Solution

Since
$$[T] = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix}$$
we see that $\operatorname{rank}([T]) = 2$, which is equal to the number of rows of $[T]$. This means that $[T]\vec{x} = \vec{y}$ is consistent for every $\vec{y} \in \mathbb{R}^2$. That is, for every $\vec{y} \in \mathbb{R}^2$, there is an $\vec{x} \in \mathbb{R}^3$ such that $T(\vec{x}) = \vec{y}$. Hence T is surjective.

Example 4

Determine whether the linear mapping $S\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$ is surjective.

Solution

Since
$$[S] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$
we see that $\operatorname{rank}([S]) = 2$ which is less than the number of rows of of $[S]$. Therefore $[S]\vec{x} = \vec{y}$ will be inconsistent for some $\vec{y} \in \mathbb{R}^3$, for example, $\vec{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. So given $\vec{y} \in \mathbb{R}^3$, we are not guaranteed to find $\vec{x} \in \mathbb{R}^2$ with $S(\vec{x}) = \vec{y}$ and therefore S is not surjective.

Exercise 2

A question appears in Mobius

Bijective Linear Mappings

From the previous sections, we noticed that a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ is injective if $\operatorname{rank}([L]) = n$ and surjective if $\operatorname{rank}([L]) = m$. Thus L is both injective and surjective if $m = \operatorname{rank}([L]) = n$, that is, $m = n$. In this case, $L : \mathbb{R}^n \to \mathbb{R}^n$ and $\text{rank}([L]) = n$, so $[L]$ is invertible and thus the linear mapping L is also invertible. We give such linear mapping a special name.

Definition
A linear mapping $L : \mathbb{R}^n \to \mathbb{R}^n$ is **bijective** if it is both injective and surjective.

A **bijective** linear mapping is one which maps each element in the domain to a unique element in the codomain and which maps on to each element of the codomain exactly once.

Remark
Note that although the concept of a bijective linear mapping relates the concepts of injective and surjective mappings, it is possible for a linear mapping to be

- injective but not surjective, or
- surjective but not injective.

Try to come up with examples of such mappings!

Example 5

Determine whether the linear mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 \\ x_1 - x_2 \end{bmatrix}$ is bijective.

Solution

Since
$$[T] = \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
we have that $\text{rank}([T]) = 2$, which means that T is both injective and surjective and, therefore, bijective.

Exercise 3

A question appears in Mobius

6.6 - Special Transformations (Rotations, Projections, Reflections)

Special Mappings: Rotations

About Rotations

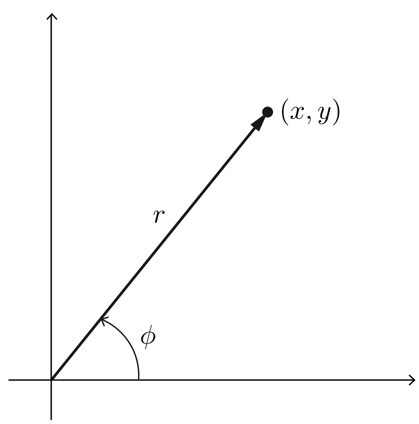
There are many types of geometrical mappings which are commonly used in a variety of applications including engineering and physics. In the following lessons, we will build an inventory of these special mappings.

We begin our study of special mappings with rotations. As we will see, rotations are linear: multiplying a vector by a scalar, then rotating it is equivalent to rotating the vector, then multiplying it by the scalar since rotations don't affect lengths.

Before we proceed to examples, we need to recall what **polar coordinates** are. The Cartesian coordinates (x, y) of a non-zero vector $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$ can be expressed in terms of the length $r = \|\vec{x}\|$ and the angle ϕ as follows:

$$\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r\cos\phi \\ r\sin\phi \end{bmatrix}$$

In the case when $\vec{x} = \vec{0}$ we have $r = 0$ and we assume that $\phi = 0$. We refer to the pair (r, ϕ) as the **polar coordinates** of $\vec{x}$. The following picture demonstrates the relation between (x, y) and (r, ϕ).



Example

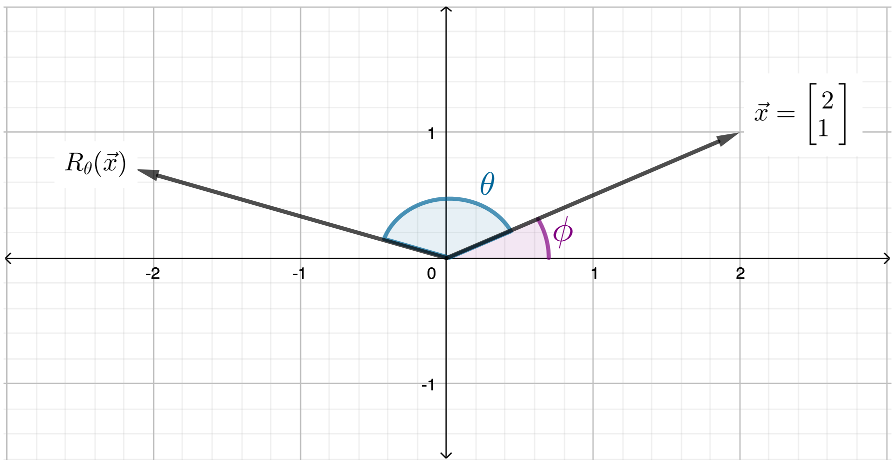
Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be a counterclockwise rotation about the origin by an angle of θ.

1. Show that R_θ is linear.
2. Compute the counterclockwise rotation of $\vec{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ about the origin by an angle of $3\pi/4$.

Solution - Part A

To see that R_θ is linear, let $\vec{x} = \begin{bmatrix} r\cos\phi \\ r\sin\phi \end{bmatrix}$, where $r = \|\vec{x}\|$ is the length of $\vec{x}$ and $\phi \in \mathbb{R}$.

Then R_θ maps the vector $\vec{x} = \begin{bmatrix} r\cos\phi \\ r\sin\phi \end{bmatrix}$ to the vector $R_\theta(\vec{x}) = \begin{bmatrix} r\cos(\phi+\theta) \\ r\sin(\phi+\theta) \end{bmatrix}$.



To prove that R_θ is linear, we will make use of the trigonometric identities for $\cos(\phi+\theta)$ and $\sin(\phi+\theta)$:

$$\begin{aligned} R_\theta(\vec{x}) &= \begin{bmatrix} r\cos(\phi+\theta) \\ r\sin(\phi+\theta) \end{bmatrix} \\ &= \begin{bmatrix} r\cos\phi\cos\theta - r\sin\phi\sin\theta \\ r\sin\phi\cos\theta + r\cos\phi\sin\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta(r\cos\phi) - \sin\theta(r\sin\phi) \\ \sin\theta(r\cos\phi) + \cos\theta(r\sin\phi) \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} r\cos\phi \\ r\sin\phi \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \vec{x} \end{aligned}$$

So R_θ is a matrix transformation and thus a linear mapping. We also see that

$$[R_\theta] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Solution - Part B

Now, for $\vec{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, we can compute the counterclockwise rotation of $\vec{x}$ about the origin by an angle of $\theta = 3\pi/4$.

$$\begin{aligned} R_{3\pi/4}(\vec{x}) &= \begin{bmatrix} \cos(3\pi/4) & -\sin(3\pi/4) \\ \sin(3\pi/4) & \cos(3\pi/4) \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -3/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \end{aligned}$$

This is also demonstrated in the picture above.

Making Connections

In this interactive example, select and drag vector $\vec{x}$ through an angle $\theta \in [0, 2\pi)$, to see how the corresponding vector $R_\theta(\vec{x})$ changes.

External resource: https://www.geogebra.org/material/iframe/id/yrxam7px/

Created with `https://www.geogebra.org`.

Exercise

A question appears in Mobius

Here are some commonly used rotation matrices in $\mathbb{R}^3$:

- $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$ gives a counterclockwise rotation by α about the x-axis.
- $\begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}$ gives a counterclockwise rotation by β about the y-axis.
- $\begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$ gives a counterclockwise rotation by γ about the z-axis.

Special Mappings: Projections

About Projections

Next, we revisit the concept of projections and view them as linear mappings.

Example

Let $\vec{d} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\vec{x}) = \text{proj}_{\vec{d}}\vec{x}$ for every $\vec{x} \in \mathbb{R}^2$. (Recall that $\text{proj}_{\vec{d}}\vec{x}$ means the projection of $\vec{x}$ onto $\vec{d}$.)

1. Show that T is linear.
2. Find the standard matrix of T.
3. Calculate $T\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix}\right)$.

Solution - Part A

To show that T is linear, we prove that it satisfies properties L1 and L2.

For property L1: Let $\vec{x}, \vec{y} \in \mathbb{R}^2$. Then,

$$\begin{aligned} T(\vec{x}+\vec{y}) &= \text{proj}_{\vec{d}}(\vec{x}+\vec{y}) \\ &= \frac{(\vec{x}+\vec{y}) \cdot \vec{d}}{\|\vec{d}\|^2}\vec{d} \\ &= \frac{\vec{x} \cdot \vec{d}}{\|\vec{d}\|^2}\vec{d} + \frac{\vec{y} \cdot \vec{d}}{\|\vec{d}\|^2}\vec{d} \\ &= \text{proj}_{\vec{d}}\vec{x} + \text{proj}_{\vec{d}}\vec{y} \\ &= T(\vec{x}) + T(\vec{y}) \end{aligned}$$

For property L2: Let $\vec{x} \in \mathbb{R}^2$ and $k \in \mathbb{R}$. Then,

$$\begin{aligned} T(k\vec{x}) &= \text{proj}_{\vec{d}}(k\vec{x}) \\ &= \frac{(k\vec{x}) \cdot \vec{d}}{\|\vec{d}\|^2}\vec{d} \\ &= k\frac{\vec{x} \cdot \vec{d}}{\|\vec{d}\|^2}\vec{d} \\ &= k\,\text{proj}_{\vec{d}}\vec{x} \\ &= kT(\vec{x}) \end{aligned}$$

Since T satisfies properties L1 and L2, then T is linear.

Solution - Part B

To find the standard matrix of T, we calculate $T(\vec{e}_1)$ and $T(\vec{e}_2)$:

$$T(\vec{e}_1) = \text{proj}_{\vec{d}}\vec{e}_1 = \frac{\vec{e}_1 \cdot \vec{d}}{\|\vec{d}\|^2}\vec{d} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

and

$$T(\vec{e}_2) = \text{proj}_{\vec{d}}\vec{e}_2 = \frac{\vec{e}_2 \cdot \vec{d}}{\|\vec{d}\|^2}\vec{d} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

so

$$[T] = [T(\vec{e}_1) \;\; T(\vec{e}_2)] = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

is the standard matrix of T.

Solution - Part C

Once we have the standard matrix of T, calculating $T(\vec{x})$ is as simple as multiplying $[T]\vec{x}$:

$$\text{proj}_{\vec{d}} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = T\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

We can generalize the previous example by letting $\vec{d}$ be an arbitrary non-zero vector in $\mathbb{R}^2$ and proceeding analogously.

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

Special Mappings: Reflections

About Reflections

Another set of special mappings are reflections. We will consider two types of reflections, namely

- reflections in $\mathbb{R}^2$ in the line passing through the origin in the direction $\vec{d}$:

$$\operatorname{refl}_{\vec{d}}(\vec{x}) = \vec{x} - 2\operatorname{perp}_{\vec{d}}\vec{x}$$

- reflection in $\mathbb{R}^3$ in the plane P:

$$\operatorname{refl}_P(\vec{x}) = \vec{x} - 2\operatorname{perp}_P \vec{x}$$

Though it is possible to generalize reflections to arbitrary vector space $\mathbb{R}^n$, we will not do that in this course.

Making Connections

Click and drag the tip of vector $\vec{x}$ and the point $D(d_1, d_2)$ to see what the reflection of the vector $\vec{x}$ in the line passing through the origin in the direction $\vec{OD}$ looks like. Notice what happens when the vector $\vec{x}$ lies on the line.

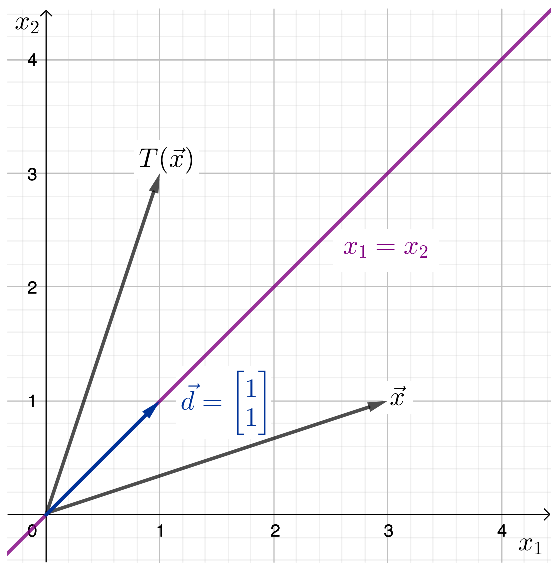
External resource: https://www.geogebra.org/material/iframe/id/we787ugh/

As the next examples show, reflections are also linear mappings.

Example 1

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a reflection in the line $x_2 = x_1$. Let $\vec{d} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then, using the definition of reflections in $\mathbb{R}^2$ (in the line passing through the origin in the direction $\vec{d}$) given at the beginning of this section, we have

$$\begin{aligned} T(\vec{x}) &= \vec{x} - 2\operatorname{perp}_{\vec{d}}\vec{x} \\ &= \vec{x} - 2(\vec{x} - \operatorname{proj}_{\vec{d}}\vec{x}) && \text{by definition of perp} \\ &= -\vec{x} + 2\operatorname{proj}_{\vec{d}}\vec{x} \end{aligned}$$



1. Show that T is linear.
2. Find its standard matrix.

Solution - Part A

To show that T is linear, we show that T satisfies properties L1 and L2.

For property L1: Let $\vec{x}, \vec{y} \in \mathbb{R}^2$. Then,

$$\begin{aligned} T(\vec{x}+\vec{y}) &= -(\vec{x}+\vec{y}) + 2\,\mathrm{proj}_{\vec{d}}(\vec{x}+\vec{y}) \\ &= -\vec{x} - \vec{y} + 2(\mathrm{proj}_{\vec{d}}\vec{x} + \mathrm{proj}_{\vec{d}}\vec{y}) && \text{since projection is linear} \\ &= -\vec{x} + 2\,\mathrm{proj}_{\vec{d}}\vec{x} - \vec{y} + 2\,\mathrm{proj}_{\vec{d}}\vec{y} \\ &= T(\vec{x}) + T(\vec{y}). \end{aligned}$$

For property L2: Let $\vec{x} \in \mathbb{R}^2$ and $k \in \mathbb{R}$. Then,

$$\begin{aligned} T(k\vec{x}) &= -k\vec{x} + 2\,\mathrm{proj}_{\vec{d}}(k\vec{x}) \\ &= -k\vec{x} + 2k\,\mathrm{proj}_{\vec{d}}\vec{x} && \text{since projection is linear} \\ &= k(-\vec{x} + 2\,\mathrm{proj}_{\vec{d}}\vec{x}) \\ &= kT(\vec{x}). \end{aligned}$$

Since T satisfies properties L1 and L2, T is linear.

Solution - Part B

Next, we find its standard matrix. We have

$$T(\vec{e}_1) = -\vec{e}_1 + 2\,\mathrm{proj}_{\vec{d}}\vec{e}_1 = -\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\left(\frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and

$$T(\vec{e}_2) = -\vec{e}_2 + 2\,\mathrm{proj}_{\vec{d}}\vec{e}_2 = -\begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2\left(\frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

So

$$[T] = [T(\vec{e}_1)\ T(\vec{e}_2)] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

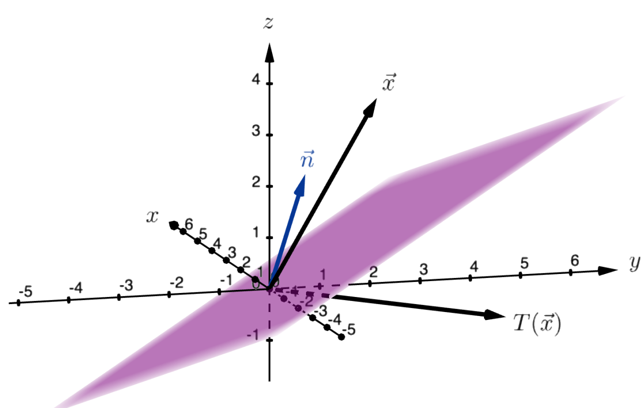
is the standard matrix for T. For example, $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$.

From here we see that a reflection in the line $x_2 = x_1$ swaps the x_1 and x_2 coordinates.

Example 2

Let $S : \mathbb{R}^3 \to \mathbb{R}^3$ be a reflection in the plane with equation $x_1 - x_2 + 2x_3 = 0$. Let $\vec{n} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$. Then,

$$S(\vec{x}) = \vec{x} - 2\operatorname{perp}_P \vec{x} = \vec{x} - 2\operatorname{proj}_{\vec{n}} \vec{x}$$



1. Show that S is linear.
2. Find its standard matrix.

Solution - Part A

To show that S is linear, we show that S satisfies properties L1 and L2.

For property L1: Let $\vec{x}, \vec{y} \in \mathbb{R}^3$. Then,

$$\begin{aligned} S(\vec{x}+\vec{y}) &= (\vec{x}+\vec{y}) - 2\operatorname{proj}_{\vec{n}}(\vec{x}+\vec{y}) \\ &= (\vec{x}+\vec{y}) - 2(\operatorname{proj}_{\vec{n}} \vec{x} + \operatorname{proj}_{\vec{n}} \vec{y}) && \text{since projection is linear} \\ &= \vec{x} - 2\operatorname{proj}_{\vec{n}} \vec{x} + \vec{y} - 2\operatorname{proj}_{\vec{n}} \vec{y} \\ &= S(\vec{x}) + S(\vec{y}) \end{aligned}$$

For property L2: Let $\vec{x} \in \mathbb{R}^3$ and $k \in \mathbb{R}$. Then,

$$\begin{aligned} S(k\vec{x}) &= k\vec{x} - 2\operatorname{proj}_{\vec{n}}(k\vec{x}) \\ &= k\vec{x} - 2k\operatorname{proj}_{\vec{n}} \vec{x} && \text{since projection is linear} \\ &= k(\vec{x} - 2\operatorname{proj}_{\vec{n}} \vec{x}) \\ &= kS(\vec{x}) \end{aligned}$$

Since S satisfies properties L1 and L2, S is linear.

Solution - Part B

Next, we find its standard matrix. We have

$$S(\vec{e}_1) = \vec{e}_1 - 2\,\mathrm{proj}_{\vec{n}}\,\vec{e}_1 = \vec{e}_1 - 2\frac{\vec{e}_1\cdot\vec{n}}{\|\vec{n}\|^2}\vec{n} = \begin{bmatrix}1\\0\\0\end{bmatrix} - 2\left(\frac{1}{6}\begin{bmatrix}1\\-1\\2\end{bmatrix}\right) = \begin{bmatrix}2/3\\1/3\\-2/3\end{bmatrix}$$

$$S(\vec{e}_2) = \vec{e}_2 - 2\,\mathrm{proj}_{\vec{n}}\,\vec{e}_2 = \vec{e}_2 - 2\frac{\vec{e}_2\cdot\vec{n}}{\|\vec{n}\|^2}\vec{n} = \begin{bmatrix}0\\1\\0\end{bmatrix} - 2\left(\frac{(-1)}{6}\begin{bmatrix}1\\-1\\2\end{bmatrix}\right) = \begin{bmatrix}1/3\\2/3\\2/3\end{bmatrix}$$

and

$$S(\vec{e}_3) = \vec{e}_3 - 2\,\mathrm{proj}_{\vec{n}}\,\vec{e}_3 = \vec{e}_3 - 2\frac{\vec{e}_1\cdot\vec{n}}{\|\vec{n}\|^2}\vec{n} = \begin{bmatrix}0\\0\\1\end{bmatrix} - 2\left(\frac{2}{6}\begin{bmatrix}1\\-1\\2\end{bmatrix}\right) = \begin{bmatrix}-2/3\\2/3\\-1/3\end{bmatrix}$$

So the standard matrix of S is

$$[S] = \begin{bmatrix}2/3 & 1/3 & -2/3\\1/3 & 2/3 & 2/3\\-2/3 & 2/3 & -1/3\end{bmatrix}$$

For example, the reflection of $\begin{bmatrix}3\\-1\\3\end{bmatrix}$ in the plane $x_1 - x_2 + 2x_3 = 0$ is given by

$$S\left(\begin{bmatrix}3\\-1\\3\end{bmatrix}\right) = \begin{bmatrix}2/3 & 1/3 & -2/3\\1/3 & 2/3 & 2/3\\-2/3 & 2/3 & -1/3\end{bmatrix}\begin{bmatrix}3\\-1\\3\end{bmatrix} = \begin{bmatrix}-1/3\\7/3\\-11/3\end{bmatrix}$$

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

Special Mappings: Stretches and Compressions

About Stretches and Compressions

Let t_1 and t_2 be fixed real numbers. The next linear mapping we consider is

$$S\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=\begin{bmatrix}t_1x_1\\t_2x_2\end{bmatrix}$$

We see that the standard matrix of S is given by

$$[S]=\begin{bmatrix}t_1 & 0\\0 & t_2\end{bmatrix}$$

Notice that, when S is applied to the vector $\vec{x}\in\mathbb{R}^2$, it scales $\vec{x}$ in the direction of $\vec{e}_1$ by a factor of t_1 and in the direction of $\vec{e}_2$ by a factor of t_2.

Making Connections

In this interactive example, choose a vector $\vec{x}$ and $t_1, t_2 \in [-3, 3]$, and see how the corresponding vector $S(\vec{x})$ changes. Notice what happens when you set $t_1 = t_2$, $t_1 = 0$ or $t_2 = 0$.

External resource: https://www.geogebra.org/material/iframe/id/dnpz8sgf/

From the previous exercise, we see that

- If $|t_i| > 1$, then a stretch in the direction of $\vec{e}_i$ occurs.
- If $0 < |t_i| < 1$, then a compression in the direction of $\vec{e}_i$ occurs.

- If $t_i = 0$, then the i-th coordinate of $S(\vec{x})$ is equal to zero.
- If $t_i = 1$, then the i-th coordinate of $S(\vec{x})$ remains unchanged.
- If $t_1 < 0$, then a reflection in the line passing through the origin in the direction of $\vec{e}_2$ occurs.
- If $t_2 < 0$, then a reflection in the line passing through the origin in the direction of $\vec{e}_1$ occurs.

Example 1

Let $A = \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}$ for $t \in \mathbb{R},\ t > 0$, and define $S(\vec{x}) = A\vec{x}$ for $\vec{x} \in \mathbb{R}^2$.

Describe the transformations of S.

Solution

We have

$$S\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} tx_1 \\ x_2 \end{bmatrix}$$

so S is a stretch when $t > 1$ or a compression when $0 < t < 1$ in the direction of $\vec{e}_1$. When $t = 0$, S is a projection onto the x_2-axis, and when $t < 0$, S is a reflection in the x_2-axis and a compression/stretch.

Example 2

Let $A = \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} \quad t \in \mathbb{R}, t > 0$ and define $T(\vec{x}) = A\vec{x}$.

Describe the transformations of T.

Solution

We have

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} tx_1 \\ tx_2 \end{bmatrix} = t\vec{x}$$

T is a contraction when $0 < t < 1$ and a stretch when $t > 1$. When $t = 1$, it does nothing. When $t < 0$, it is a reflection through the origin along with a contraction/stretch.

Exercise

A question appears in Mobius

Unit 7

Determinants

7.1 - Finding Determinants via Cofactor Expansion

Determinants and Adjugates for 2×2 Matrices

Let's start with some basic definitions.

Definitions
Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix. The **determinant** of A is

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The **adjugate** of A is

$$\operatorname{adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

As we will see, the **determinant** and the **adjugate** of a 2×2 matrix A are closely related to its inverse, A^{-1}. The following example gives us some intuition.

Example

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

1. Calculate the determinant of A.
2. Calculate the adjugate of A.

Solution - Part A

Using the formula for the determinant, we find

$$\det(A) = 1(4) - 2(3) = 4 - 6 = -2$$

Solution - Part B

The adjugate of A is

$$\operatorname{adj}(A) = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$$

Notice that, to find the adjugate, we swapped the entries on the diagonal of A and put a negative sign in front of the off-diagonal entries.

We also make the following observations:

$$A(\operatorname{adj}(A)) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = -2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \det(A)I$$

and

$$(\operatorname{adj}(A))A = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = -2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \det(A)I$$

From this we see that

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \left(\frac{1}{-2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

that is, $A\left(\dfrac{1}{\det(A)} \operatorname{adj}(A)\right) = I$, which means that

$$A^{-1} = \left(\frac{1}{\det(A)} \operatorname{adj}(A)\right) = \left(\frac{1}{-2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}\right) = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & \frac{-1}{2} \end{bmatrix}$$

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

The observation made in the previous example can be generalized as follows.

Theorem 1: Invertibility and the Determinant
Let A be a 2×2 matrix. Then,

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

Moreover, A is invertible if and only if $\det(A) \neq 0$ and in this case

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

Proof

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix.

Then, by definition, $\det(A) = ad - bc$ and $\mathrm{adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Now

$$A(\mathrm{adj}(A)) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = (\det(A))I$$

and

$$(\mathrm{adj}(A))A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = (\det(A))I$$

To prove the second part of the theorem, we will show that 1) $\det(A) \neq 0$ implies that A is invertible; and 2) that A being invertible implies that $\det(A) \neq 0$.

Assume that $\det(A) \neq 0$. From

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

We obtain

$$A\left(\frac{1}{\det(A)}\,\mathrm{adj}(A)\right) = I = \left(\frac{1}{\det(A)}\,\mathrm{adj}(A)\right)A$$

So

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

Thus $\det(A) \neq 0$, which means that A is invertible and gives our formula for A^{-1}.

We now show A is invertible implies $\det(A) \neq 0$.

Assume for a contradiction that $\det(A) = 0$. If A is the zero matrix then we reach a contradiction immediately, because for any 2×2 matrix B the matrices AB and BA are equal to the zero matrix (as opposed to the identity matrix). Thus we may assume that A is not equal to the zero matrix.

Since A is invertible and it differs from the zero matrix, at least one of a, b, c, and d are not zero. Since

$$A(\operatorname{adj}(A)) = \det(A)I = 0I = 0$$

We have

$$A \begin{bmatrix} d \\ -c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} -b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Because either

$$\begin{bmatrix} d \\ -c \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} -b \\ a \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We see that the homogeneous system $A\vec{x} = \vec{0}$ has a nontrivial solution, so A is not invertible by the Invertible Matrix Theorem. This is a contradiction, so $\det(A) \neq 0$.

□

Remark
The name **determinant** is no accident: the determinant "determines" whether a matrix A is invertible or not!

Exercise 3

A question appears in Mobius

Determinants and Adjugates for $n \times n$ Matrices

Calculating determinants for $n \times n$ matrices is not as straightforward as it is for 2×2 matrices. In order to calculate these determinants, we will need some additonal tools.

Definition
Let A be a real $n \times n$ matrix and let A_{ij} be the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i^{th} row and j^{th} column of A. The (i, j)-**cofactor** of A, denoted by C_{ij}, is

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

Remark
Note an important difference in notation: a_{ij} denotes a **number**, while A_{ij} denotes a **matrix**.

Example 1

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

1. Find the $(3, 2)$-cofactor of A.
2. Find the $(2, 2)$-cofactor of A.

Solution - Part A

Using the definition of the cofactor, we have

$$\begin{aligned} C_{32} &= (-1)^{3+2} \det(A_{32}) \\ &= (-1)^5 \begin{vmatrix} 1 & \cancel{-2} & 3 \\ 1 & \cancel{0} & 4 \\ \cancel{4} & \cancel{1} & \cancel{4} \end{vmatrix} \\ &= (-1)^5 \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} \\ &= (-1)(4-3) \\ &= -1 \end{aligned}$$

Therefore, -1 is the $(3, 2)$-cofactor of A.

Solution - Part B

Again using the definition of the cofactor, we have

$$\begin{aligned}
C_{22} &= (-1)^{2+2}\det(A_{22}) \\
&= (-1)^4 \begin{vmatrix} 1 & \cancel{-2} & 3 \\ \cancel{1} & \cancel{0} & \cancel{4} \\ 4 & \cancel{1} & 4 \end{vmatrix} \\
&= (-1)^4 \begin{vmatrix} 1 & 3 \\ 4 & 4 \end{vmatrix} \\
&= 1(4-12) \\
&= -8
\end{aligned}$$

Therefore, -8 is the $(2,2)$-cofactor of A.

Exercise 1

A question appears in Mobius

We can use cofactors to define determinants for $n \times n$ matrices.

Definition
Let A be an $n \times n$ matrix. The **determinant** of A is given by

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

which is a cofactor expansion of A along the i^{th} row of A.
Equivalently, we can define

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

which is a cofactor expansion of A along the j^{th} column of A.

Remark
The definition of the determinant allows us to do cofactor expansion along any row or column we like. In practice, we try to choose rows or columns that have zeros in them to simplify our calculations.

Example 2

Given $A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 3 & 4 \\ 5 & 6 & -7 \end{bmatrix}$, compute $\det(A)$.

Solution

Expanding along the first column,

$$\begin{aligned} \det(A) &= a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} \\ &= 1(-1)^{1+1}\begin{vmatrix} 3 & 4 \\ 6 & -7 \end{vmatrix} + 0(-1)^{2+1}\begin{vmatrix} 0 & -2 \\ 6 & -7 \end{vmatrix} + 5(-1)^{3+1}\begin{vmatrix} 0 & -2 \\ 3 & 4 \end{vmatrix} \\ &= 1(-21-24) + 5(0+6) \\ &= -45 + 30 \\ &= -15 \end{aligned}$$

Notice that when we expanded along the first column, the zero entry at a_{21} simplified our calculations.

Exercise 2

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & -5 & 6 \\ -7 & 8 & 9 \end{bmatrix}$. Perform the cofactor expansion along the first row.

A question appears in Mobius

Exercise 3

A question appears in Mobius

Let's calculate the determinant of a 4×4 matrix.

Example 3

Given $A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 3 \\ -1 & 1 & 2 & 1 \end{bmatrix}$, find $\det(A)$.

Solution

We notice that the third row has lots of zeros in it. We will therefore use cofactor expansion along the third row:

$$\begin{aligned}
\det(A) &= \begin{vmatrix} 1 & 2 & -1 & 3 \\ 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 3 \\ -1 & 1 & 2 & 1 \end{vmatrix} && \text{expand along the third row} \\
&= -3 \begin{vmatrix} 1 & 2 & -1 \\ 1 & 2 & 0 \\ -1 & 1 & 2 \end{vmatrix} && \text{now expand along the third column} \\
&= -3 \left(-1 \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} + 2 \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} \right) \\
&= -3(-1(1+2) + 2(2-2)) \\
&= -3(-3+0) \\
&= 9
\end{aligned}$$

Exercise 4

A question appears in Mobius

Definitions
Let $A = [a_{ij}]$ be a real $n \times n$ matrix,

- $C_{ij} = (-1)^{i+j} \det(A_{ij})$ is the (i, j)**-cofactor** of A
- The **cofactor matrix** of A is
$$\mathrm{Cof}(A) = [C_{ij}]$$
- The **adjugate** of A is
$$\mathrm{adj}(A) = [C_{ij}]^T$$

Note that, in this lesson, we have introduced two definitions of the **adjugate** to demonstrate two different methods for calculating it.

Example 4

Given $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix}$, find $\mathrm{adj}(A)$.

Solution

$$\begin{aligned}
\operatorname{adj}(A) &= \begin{bmatrix} \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 3 & 5 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 3 & 4 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 3 & 5 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} \end{bmatrix}^T \\
&= \begin{bmatrix} -3 & 1 & 1 \\ 2 & -4 & 2 \\ 1 & 1 & -1 \end{bmatrix}^T \\
&= \begin{bmatrix} -3 & 2 & 1 \\ 1 & -4 & 1 \\ 1 & 2 & -1 \end{bmatrix}
\end{aligned}$$

Now, notice that we can easily compute $\det(A)$

$$\det(A) = 1(-3) + 2(1) + 3(1) = 2$$

which is a cofactor expansion along the first row of A. Also note that this is the dot product of the first row of A and the first column of $\operatorname{adj}(A)$:

$$\begin{aligned}
A(\operatorname{adj}(A)) &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} -3 & 2 & 1 \\ 1 & -4 & 1 \\ 1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \\
&= 2I \\
&= (\det(A))I
\end{aligned}$$

You can also check that $(\operatorname{adj}(\mathrm{A}))A = 2I$.

As the next theorem shows, we can use both the determinant of A and the adjugate of A to determine A^{-1}.

Theorem 2
Let A be an $n \times n$ matrix. Then

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

Moreover, A is invertible if and only if $\det(A) \neq 0$. In this case,

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

Example 5

Let $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 4 \\ 1 & 2 & 4 \end{bmatrix}$.

1. Find $\det(A)$.
2. Find $\operatorname{adj}(A)$.
3. Find A^{-1}.

Solution - Part A

To find $\det(A)$, let's use a cofactor expansion along the first row:

$$\begin{aligned}\det(A) &= 1\begin{vmatrix} 1 & 4 \\ 2 & 4 \end{vmatrix} - 1\begin{vmatrix} 1 & 4 \\ 1 & 4 \end{vmatrix} + 2\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} \\ &= 1(4-8) - 1(4-4) + 2(2-1) \\ &= -4+2 \\ &= -2\end{aligned}$$

Solution - Part B

Now find $\operatorname{adj}(A)$:

$$\begin{aligned}\operatorname{adj}(A) &= \begin{bmatrix} \begin{vmatrix} 1 & 4 \\ 2 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 4 \\ 1 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} \\ -\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} \\ \begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} \end{bmatrix}^T \\ &= \begin{bmatrix} -4 & 0 & 1 \\ 0 & 2 & -1 \\ 2 & -2 & 0 \end{bmatrix}^T \\ &= \begin{bmatrix} -4 & 0 & 2 \\ 0 & 2 & -2 \\ 1 & -1 & 0 \end{bmatrix}\end{aligned}$$

Solution - Part C

Lastly, we find A^{-1}:

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} -4 & 0 & 2 \\ 0 & 2 & -2 \\ 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -1 \\ 0 & -1 & 1 \\ \frac{-1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

Exercise 5

A question appears in Mobius

7.2 - Finding Determinants via Row Reduction

Finding Determinants via Row Reduction

There are other ways of calculating determinants in addition to using cofactor expansion. In this lesson, we will see how to use matrix row reduction to find the determinant of a matrix.

Elementary Row/Column Operations

When we row reduce a matrix, we perform elementary operations on its rows and/or columns. Each of these operations has a particular relationship with the determinant of the matrix. In fact, the determinant changes predictably under elementary row and column operations. Let's explain these relationships by some examples.

Example 1

Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}$. Determine the effects of elementary row operations on $\det(A)$; namely:

1. Swapping rows.
2. Adding a multiple of one row to another row.
3. Multiplying a row by a non-zero number.

Solution - Part A

We have $\det(A) = (1)(4) - (1)(2) = 2$.

If we swap the rows of A, we get

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \underset{R_1 \leftrightarrow R_2}{\sim} \begin{bmatrix} 1 & 4 \\ 1 & 2 \end{bmatrix}$$

The determinant of the new matrix $B = \begin{bmatrix} 1 & 4 \\ 1 & 2 \end{bmatrix}$ is $\det(B) = (1)(2) - (4)(1) = -2 = -\det(A)$.

Solution - Part B

If we add R_1 to R_2 of A, we get

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \underset{R_1+R_2}{\sim} \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}$$

The determinant of the new matrix $C = \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}$ is $\det(C) = (1)(6) - (2)(2) = 2 = \det(A)$.

Solution - Part C

If we multiply R_1 of A by 2, we get

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \underset{2R_1}{\sim} \begin{bmatrix} 2 & 4 \\ 1 & 4 \end{bmatrix}$$

The determinant of the new matrix $D = \begin{bmatrix} 2 & 4 \\ 1 & 4 \end{bmatrix}$ is $\det(D) = (2)(4) - (1)(4) = 4 = 2\det(A)$.

Example 2

Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}$. Determine the effects of elementary column operations on $\det(A)$; namely:

1. Swapping columns.
2. Adding a multiple of one column to another column.
3. Multiplying a column by a non-zero number.

Solution - Part A

We have $\det(A) = (1)(4) - (1)(2) = 2$.

If we swap the columns of A, we get

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \underset{C_1 \leftrightarrow C_2}{\sim} \begin{bmatrix} 2 & 1 \\ 4 & 1 \end{bmatrix}$$

The determinant of the new matrix $B = \begin{bmatrix} 2 & 1 \\ 4 & 1 \end{bmatrix}$ is $\det(B) = (2)(1) - (1)(4) = -2 = -\det(A)$.

Solution - Part B

If we add C_1 to C_2 of A, we get

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \underset{C_2+C_1}{\sim} \begin{bmatrix} 1 & 3 \\ 1 & 5 \end{bmatrix}$$

The determinant of the new matrix $C = \begin{bmatrix} 1 & 3 \\ 1 & 5 \end{bmatrix}$ is $\det(C) = (1)(5) - (1)(3) = 2 = \det(A)$.

Solution - Part C

If we multiply C_1 of A by 2, we get

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \underset{2C_1}{\sim} \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

The determinant of the new matrix $D = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$ is $\det(D) = (2)(4) - (2)(2) = 4 = 2\det(A)$.

Let's summarize what we observed in the previous two examples, as well as a few more characteristics of determinant and row/column operations.

Theorem 1: Determinant and Row/Column Operations
Let A be an $n \times n$ matrix.

1. If A has a row (or column) of zeros, then $\det(A) = 0$.
2. If B is obtained from A by swapping two distinct rows (or two distinct columns), then $\det(B) = -\det(A)$.
3. If B is obtained from A by adding a multiple of one row to another row (or a multiple of one column to another column), then $\det(A) = \det(B)$.
4. If two distinct rows of A (or two distinct columns of A) are equal, then $\det(A) = 0$.
5. If B is obtained from A by multiplying a row (or a column) by $k \in \mathbb{R}$, then $\det(B) = k\det(A)$.

Making Connections

Consider the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, whose determinant is $\det(A) = (1)(4) - (2)(3) = -2$.

Press the buttons to obtain a matrix B which is row equivalent to A, and observe how the determinant of B changes.

External resource: https://www.geogebra.org/material/iframe/id/udyqazqa/

Conducting these same operations on R_2 , or on columns instead of rows, will produce similar results, as per the theorem introduced earlier.

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

Using the Theorem to find the Determinant

We often use the results of the theorem on determinants and elementary row/column operations to help us calculate the determinant of a matrix. Let's see how they are used in practice.

Example 3

Find $\det(A)$ if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix}$$

Solution

We start by applying some elementary row operations in order to make cofactor expansion simpler: we will make the first column $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ using elementary row operations, then do cofactor expansion along the first column.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix} \underset{R_2 - 4R_1}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 7 & 8 & 10 \end{bmatrix} \underset{R_3 - 7R_1}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{bmatrix}$$

Note that neither of the elementary row operations performed affects the determinant of A since we added a multiple of R_1 to R_2 and a multiple of R_1 to R_3.

Now that we have a simpler matrix, we can use cofactor expansion along the first column to calculate $\det(A)$:

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{vmatrix} \\ &= 1 \begin{vmatrix} -3 & -6 \\ -6 & -11 \end{vmatrix} \\ &= (-3)(-11) - (-6)(-6) \\ &= -3 \end{aligned}$$

Remark
When using elementary row and column operations to calculate the determinant of a matrix, do not do row and column operations on the same step: each step should contain only row operations or only column operations.

Example 4

Let

$$A = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}$$

Show that $\det(A) = (b-a)(c-a)(c-b)$.

Solution

We start by applying some elementary row operations in order to make cofactor expansion simpler. More precisely, we will make the first column $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ using elementary row operations, then do cofactor expansion along the first column.

$$
\begin{aligned}
\det(A) = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} &\underset{\substack{R_2-R_1 \\ R_3-R_1}}{=} \begin{vmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{vmatrix} \\
&= 1 \begin{vmatrix} (b-a) & (b-a)(b+a) \\ (c-a) & (c-a)(c+a) \end{vmatrix} \\
&= (b-a)(c-a)(c+a) - (c-a)(b-a)(b+a) \\
&= (b-a)(c-a)(c+a-b-a) \\
&= (b-a)(c-a)(c-b)
\end{aligned}
$$

Example 5

Find $\det(A)$ if

$$A = \begin{bmatrix} x & x & 1 \\ x & 1 & x \\ 1 & x & x \end{bmatrix}$$

When does $\det(A) = 0$?

Solution

We start by applying some elementary row operations in order to make cofactor expansion simpler. More precisely, we will make the first column $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ using elementary row operations, then do cofactor expansion along the first column.

$$
\begin{aligned}
\det(A) = \begin{vmatrix} x & x & 1 \\ x & 1 & x \\ 1 & x & x \end{vmatrix} &\underset{\substack{R_1-xR_3 \\ R_2-xR_3}}{=} \begin{vmatrix} 0 & x-x^2 & 1-x^2 \\ 0 & 1-x^2 & x-x^2 \\ 1 & x & x \end{vmatrix} \\
&= 1 \begin{vmatrix} x(1-x) & (1+x)(1-x) \\ (1+x)(1-x) & x(1-x) \end{vmatrix} \\
&= x^2(1-x)^2 - (1+x)^2(1-x)^2 \\
&= (1-x)^2(x^2-(1+x)^2) \\
&= (1-x)^2(x^2-1-2x-x^2) \\
&= -(1-x)^2(1+2x)
\end{aligned}
$$

So

$$\det(A) = 0 \Leftrightarrow -(1-x)^2(1+2x) = 0 \Leftrightarrow x = 1 \quad \text{or} \quad x = \frac{-1}{2}$$

Exercise 3

A question appears in Mobius

Theorem 2
If A is an $n \times n$ matrix and $k \in \mathbb{R}$, then

$$\det(kA) = k^n \det(A)$$

Note: The above Theorem is clearly true for $k = 0$. For $k \neq 0$, we may multiply each row of kA by $\frac{1}{k}$ (or factor a k out of each row in the determinant of kA) which gives the exponent of n of k.

Exercise 4

A question appears in Mobius

Determinant of Triangular Matrices

The determinant of a triangular matrix is, as we will see, particularly easy to calculate.

Example

Compute $\det(A)$ if

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 10 \end{bmatrix}$$

Solution

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 10 \end{vmatrix} \\ &= 1 \begin{vmatrix} 3 & 0 & 0 \\ 5 & 6 & 0 \\ 8 & 9 & 10 \end{vmatrix} \\ &= 1(3) \begin{vmatrix} 6 & 0 \\ 9 & 10 \end{vmatrix} \\ &= 1(3)(6)(10) \\ &= 180 \end{aligned}$$

Note that $\det(A)$ is just the product of the entries on the main diagonal of A.

This result is true for any triangular matrix:

Theorem 3: Determinant of Triangular Matrices
If $A = [a_{ij}]$ is an $n \times n$ triangular (upper or lower triangular) matrix, then

$$\det(A) = a_{11}a_{22}\dots a_{nn} = \prod_{i=1}^{n} a_{ii}$$

Remark
Notice that, for any integer n, the $n \times n$ identity matrix I is an example of a matrix that is both upper and lower triangular. Since I has only 1's on its diagonal, it follows from Theorem 3 that $\det(I) = 1$.

Exercise

A question appears in Mobius

7.3 - Properties of Determinants

Properties of Determinants

In this lesson, we will study some useful properties of determinants with respect to matrix operations. We start by examining the product of two matrices and its relation to the determinant.

Determinant and Product of Matrices

Example 1

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$.

Find both $\det(A)\det(B)$ and $\det(AB)$.

Solution

We have

$$\det(A)\det(B) = (4-6)(2-(-1)) = -2(3) = -6$$

We also have

$$\det(AB) = \begin{vmatrix} -1 & 5 \\ -1 & 11 \end{vmatrix} = -11-(-5) = -6$$

Notice that $\det(A)\det(B) = \det(AB)$. This is true in general.

Theorem 1: Determinant of a Product
For $n \times n$ matrices A_1, $A_2, \cdots, A_k$, we have

$$\det(A_1 A_2 \cdots A_k) = \det(A_1)\,\det(A_2)\cdots\det(A_k)$$

That is, the determinant of the product is the product of the determinants.
In particular, if $A_1 = A_2 = \ldots = A_k$, then we obtain

$$\det(A^k) = (\det(A))^k$$

for any positive integer k.

Exercise 1

A question appears in Mobius

Determinant of the Inverse Matrix

Next, let's examine the relationship between $\det(A)$ and $\det(A^{-1})$.

Theorem 2: Determinant of the Inverse Matrix
Let A be an invertible $n \times n$ matrix. Then

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Proof

We have $A\,A^{-1} = I$. Taking determinants gives

$$\det(A\,A^{-1}) = \det(I)$$

Using the result about the determinant of a matrix product, we have

$$\det(A)\,\det(A^{-1}) = 1$$

Since A is invertible, $\det(A) \neq 0$, we can rearrange, to find

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

□

Example 2

Let $A = \begin{bmatrix} 2 & -2 \\ 3 & 1 \end{bmatrix}$. Find $\det(A^{-1})$.

Solution

We have $\det(A) = (2)(1) - (-2)(3) = 2 + 6 = 8$.

By the theorem on the determinant of the inverse matrix,

$$\det(A^{-1}) = \frac{1}{\det(A)} = \frac{1}{8}$$

Exercise 2

A question appears in Mobius

We can use the result of the theorem to make a further observation: since $\dfrac{1}{\det(A)} = (\det(A))^{-1}$ for an invertible $n \times n$ matrix A , we have

$$\det(A^{-1}) = (\det(A))^{-1}$$

Recalling that $A^{-k} = (A^{-1})^k$ for a positive integer k,

$$\det(A^{-k}) = \det((A^{-1})^k) = (\det(A^{-1}))^k = (\det(A))^{-k}$$

Thus

$$\det(A^k) = (\det(A))^k$$

for any integer k. When k is positive, the statement is always true; however, when k is negative, the statement only holds when A is also invertible.

For the case where $k = 0$, we have $\det(A^0) = (\det(A))^0 = 1$ provided that A is invertible. By convention, A^0 is the identity matrix of the appropriate size.

Example 3

Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$. Determine $\det(A^{20})$.

Solution

We have $\det(A) = (1)(3) - (1)(1) = 3 - 1 = 2$. Using the result above, we find

$$\det(A^{20}) = (\det(A))^{20} = 2^{20}$$

Exercise 3

A question appears in Mobius

Determinant of a Matrix Transpose

Theorem 3: Determinant of a Matrix Tranpose
Let A be a real $n \times n$ matrix. Then

$$\det(A^T) = \det(A)$$

Remark
Since $\det(A^T) = \det(A)$, we see why we can do column operations to A when computing $\det(A)$: column operations on A are just row operations on A^T.

Example 4

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Calculate $\det(A^T)$.

Solution

We have $\det(A) = (1)(4) - (2)(3) = 4 - 6 = -2$.

Using the theorem about the determinant of the transpose of a matrix, we have

$$\det(A^T) = \det(A) = -2$$

Summary of Properties of the Determinant

Let's summarize all of the properties of the determinant that we've seen so far:

- $\det(A_1 A_2 \cdots A_k) = \det(A_1)\, \det(A_2) \cdots \det(A_k)$
- $\det(A^{-1}) = \dfrac{1}{\det(A)}$, provided that A is invertible
- $\det(A^k) = (\det(A))^k$ for any integer $k > 1$
- $\det(A^k) = (\det(A))^k$ for any integer $k < 0$, provided that A is invertible
- $\det(A^T) = \det(A)$

We can use these properties in combination, as shown in the example below.

Example 5

If $\det(A) = 3$, $\det(B) = -2$ and $\det(C) = 4$ for invertible $n \times n$ matrices A, B and C, find $\det\left(A^2\, B^T\, C^{-1}\, B^2 (A^{-1})^2\right)$.

Solution

We have

$$\begin{aligned}
\det\left(A^2\, B^T\, C^{-1}\, B^2 (A^{-1})^2\right) &= \det(A^2)\, \det(B^T)\, \det(C^{-1})\, \det(B^2)\, \det((A^{-1})^2) \\
&= (\det(A))^2\, \det(B)\, \frac{1}{\det(C)}\, (\det(B))^2\, \frac{1}{(\det(A))^2} \\
&= \frac{(\det(B))^3}{\det(C)} \\
&= \frac{(-2)^3}{4} \\
&= \frac{-8}{4} \\
&= -2
\end{aligned}$$

Exercise 4

A question appears in Mobius

7.4 - Area and Volume

Area of a Parallelogram

Formula for the Area of a Parallelogram

The determinant can be used to compute areas in $\mathbb{R}^2$, volumes in $\mathbb{R}^3$, and n-volumes in $\mathbb{R}^n$ for $n \geq 4$. We will start by focusing on the area calculation.

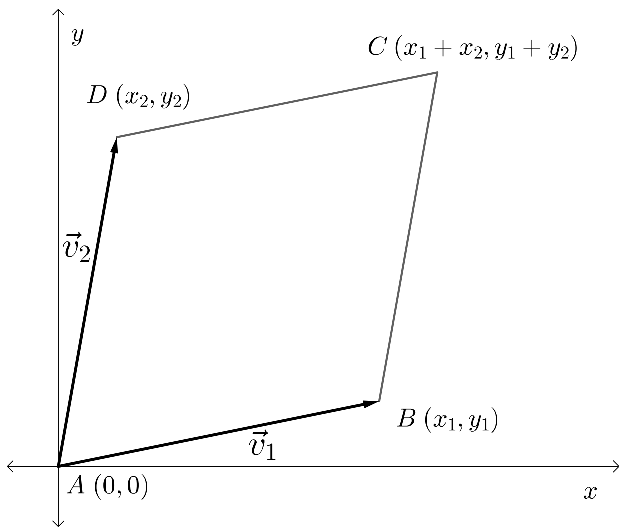
Recall that the parallelogram induced by the two vectors

$$\vec{v}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

in $\mathbb{R}^2$ is the parallelogram $ABCD$ with

$$A = (0,0), \quad B = (x_1, y_1), \quad C = (x_1 + x_2, y_1 + y_2), \quad \text{and} \quad D = (x_2, y_2).$$

This parallelogram is shown in the diagram below.



As the following proposition suggests, the determinant can be used to compute the area of the parallelogram $ABCD$.

Proposition 1
Let

$$\vec{v}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

be two vectors in $\mathbb{R}^2$. Then the area of the parallelogram induced by the vectors $\vec{v}_1$ and $\vec{v}_2$ is

$$\begin{aligned} \text{Area}(ABCD) &= \|\vec{v}_1\| \cdot \|\vec{v}_2\| \cdot |\sin\theta| \\ &= \left|\det \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix}\right| \\ &= \left|\det \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix}\right| \\ &= |x_1 y_2 - x_2 y_1| \end{aligned}$$

Remark
This means that the area of the parallelogram induced by the vectors $\vec{v}_1$ and $\vec{v}_2$ is equal to the absolute value of the determinant of the matrix whose columns are $\vec{v}_1$ and $\vec{v}_2$.

Example

Consider the area of the parallelogram $ABCD$ induced by the vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

Solution

The area of $ABCD$ is given by

$$\begin{aligned}\text{Area}(ABCD) &= \left|\det\begin{bmatrix}\vec{v}_1 & \vec{v}_2\end{bmatrix}\right| \\ &= \left|\det\begin{bmatrix}1 & 4\\ 2 & 1\end{bmatrix}\right| \\ &= |1\cdot 1 - 2\cdot 4| \\ &= |-7| \\ &= 7\end{aligned}$$

Notice the importance of absolute value: without it, the area would be negative, which is impossible. Later, we will see that the sign of the determinant $\det\begin{bmatrix}\vec{v}_1 & \vec{v}_2\end{bmatrix}$ also has a clear geometric meaning.

Exercise 1

1. Construct the parallelogram induced by the vectors $\vec{u}$ and $\vec{v}$, by dragging the two dots at (0,0) to the appropriate location in the xy-plane. The feedback indicates whether your construction is correct or incorrect.
2. Once you have the correct construction, calculate and enter the area of the parallelogram.

External resource: https://www.geogebra.org/material/iframe/id/adjtbvan/

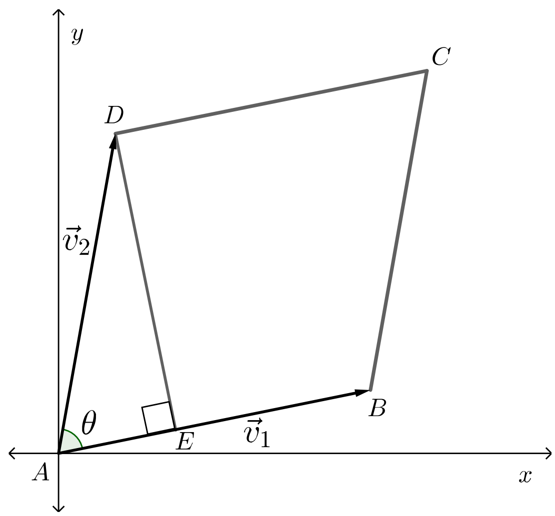
Exercise 2

A question appears in Mobius

Proof of the Area Formula

Let us now prove the formula for the area computation.

Recall that the area of a parallelogram can be computed by multiplying the length of the base by the height. Therefore, let $ABCD$ be our parallelogram with height DE and let θ denote the angle $\angle BAD$. For simplicity, we will restrict our attention to the case when $0 < \theta \leq \pi/2$, as the other cases can be established analogously.



$$\text{Area}(ABCD) = (\text{length of } AB) \cdot (\text{length of } DE)$$

Note that the length of AB is $\|\vec{v}_1\|$.

We find the height DE by considering the right triangle ADE. Since $\theta = \angle BAD = \angle EAD$,

$$\sin\theta = \frac{\text{length of } DE}{\text{length of } AD}$$

Notice that the length of AD is $\|\vec{v}_2\|$. After rearranging, we get

$$\text{length of } DE = \|\vec{v}_2\| \cdot \sin\theta$$

Hence the area of the parallelogram is

$$\begin{aligned}\text{Area}(ABCD) &= (\text{length of } AB) \cdot (\text{length of } DE)\\ &= \|\vec{v}_1\| \cdot \|\vec{v}_2\| \cdot \sin\theta\end{aligned}$$

Recall that

$$\cos\theta = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|\|\vec{v}_2\|}$$

and

$$\sin^2\theta + \cos^2\theta = 1$$

Hence, we can compute the square of the area as follows:

$$\begin{aligned}\text{Area}(ABCD)^2 &= \|\vec{v}_1\|^2 \cdot \|\vec{v}_2\|^2 \cdot \sin^2\theta\\ &= \|\vec{v}_1\|^2 \cdot \|\vec{v}_2\|^2 \cdot (1 - \cos^2\theta)\\ &= \|\vec{v}_1\|^2 \cdot \|\vec{v}_2\|^2 \cdot \left(1 - \frac{(\vec{v}_1 \cdot \vec{v}_2)^2}{\|\vec{v}_1\|^2\|\vec{v}_2\|^2}\right)\\ &= \|\vec{v}_1\|^2 \cdot \|\vec{v}_2\|^2 - (\vec{v}_1 \cdot \vec{v}_2)^2\end{aligned}$$

We then use the formulas

$$\|\vec{v}_1\|^2 = x_1^2 + y_1^2, \qquad \|\vec{v}_2\|^2 = x_2^2 + y_2^2, \qquad \text{and} \qquad \vec{v}_1 \cdot \vec{v}_2 = x_1x_2 + y_1y_2$$

to compute

$$\begin{aligned}
\text{Area}(ABCD)^2 &= \|\vec{v}_1\|^2 \cdot \|\vec{v}_2\|^2 - (\vec{v}_1 \cdot \vec{v}_2)^2 \\
&= (x_1^2 + y_1^2)(x_2^2 + y_2^2) - (x_1x_2 + y_1y_2)^2 \\
&= (x_1y_2 - y_1x_2)^2 \\
&= \det \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix}^2
\end{aligned}$$

Taking the square root of both sides of the above equation, we obtain

$$\text{Area}(ABCD) = \left| \det \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} \right|$$

□

Remark
More generally, if we allow the angle $\theta = \angle BAD$ to be arbitrary (with the convention that $\theta = 0$ when at least one of the vectors is equal to $\vec{0}$), then the area of parallelogram $ABCD$ can be computed as follows:

$$\text{Area}(ABCD) = ||\vec{v}_1|| \cdot ||\vec{v}_2|| \cdot |\sin\theta|$$

Making Connections

Observe how the area of the parallelogram induced by $\vec{u}$ and $\vec{v}$ changes with respect to other related values as you drag the points P and Q to different locations.

External resource: https://www.geogebra.org/material/iframe/id/xr6wnqdx/

Exercise 3

Drag the points P and Q so that the area of the parallelogram induced by $\vec{u}_1$ and $\vec{u}_2$ matches the required area.

External resource: https://www.geogebra.org/material/iframe/id/rnxyvsg6/

Exercise 4

A question appears in Mobius

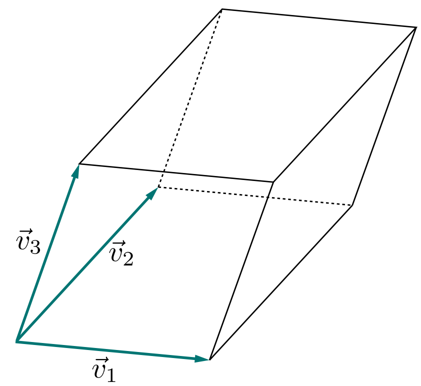
Exercise 5

A question appears in Mobius

Volume of a Parallelepiped

Formula for the Volume of a Parallelepiped

We will now turn our attention to the computation of the volume of the parallelepiped induced by three vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{R}^3$.



Proposition 2
Let

$$\vec{v}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix}$$

be vectors in $\mathbb{R}^3$. Then the volume V of the parallelepiped induced by the vectors $\vec{v}_1, \vec{v}_2$ and $\vec{v}_3$ is equal to

$$\begin{aligned} V &= |(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3| \\ &= \left|\det \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}\right| = \left|\det \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}\right| \\ &= |x_1 y_2 z_3 + x_3 y_1 z_2 + x_2 y_3 z_1 - x_1 y_3 z_2 - x_2 y_1 z_3 - x_3 y_2 z_1| \end{aligned}$$

Remark
The expression $(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3$ is called the **scalar triple product** of the vectors $\vec{v}_1, \vec{v}_2$ and $\vec{v}_3$.

Example 1

Compute the volume of the parallelepiped induced by the vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$$

Solution

The volume V of the parallelepiped can be computed as follows:

$$\begin{aligned} V &= \left| \det \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} \right| \\ &= \left| \det \begin{bmatrix} 1 & 3 & 2 \\ 2 & 1 & 2 \\ 0 & 0 & 4 \end{bmatrix} \right| \end{aligned}$$

To compute the determinant, we apply cofactor expansion across the third row:

$$\begin{aligned} V &= \left| \det \begin{bmatrix} 1 & 3 & 2 \\ 2 & 1 & 2 \\ 0 & 0 & 4 \end{bmatrix} \right| \\ &= \left| 0 \cdot \det \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} - 0 \cdot \det \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix} + 4 \cdot \det \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \right| \\ &= \left| 4 \cdot \det \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \right| \\ &= |4 \cdot (1 \cdot 1 - 2 \cdot 3)| \\ &= |-20| \\ &= 20 \end{aligned}$$

We conclude that the volume of the parallelepiped is 20.

Example 2

Compute the volume of the parallelepiped induced by the vectors

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 6 \end{bmatrix}$$

Solution

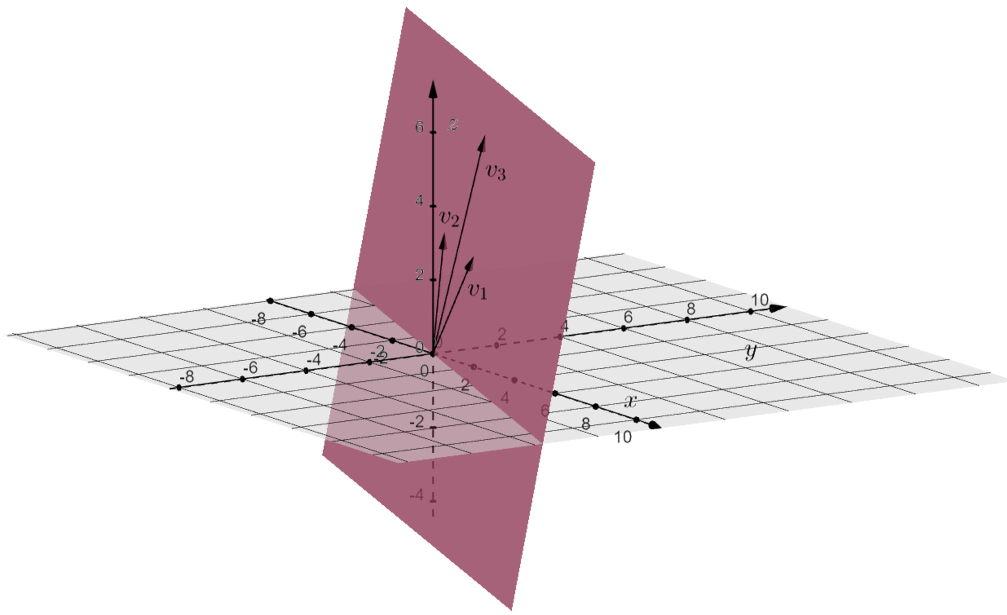
The volume V of the parallelepiped can be computed as follows:

$$\begin{aligned} V &= \left| \det \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} \right| \\ &= \left| \det \begin{bmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ 3 & 3 & 6 \end{bmatrix} \right| \end{aligned}$$

Applying cofactor expansion across the first column, we have:

$$\begin{aligned} V &= \left| 2 \cdot \det \begin{bmatrix} 1 & 1 \\ 3 & 6 \end{bmatrix} + 3 \cdot \det \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \right| \\ &= |2(1 \cdot 6 - 3 \cdot 1) + 3(-1 \cdot 1 - 1 \cdot 1)| \\ &= 0 \end{aligned}$$

This result makes sense, because these three vectors lie in the same plane. This is shown in the diagram below.



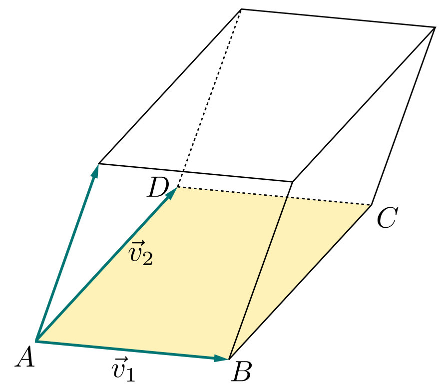
Exercise 1

A question appears in Mobius

Proof of the Volume Formula

We will now prove the formula for the volume computation. Recall that the volume of a parallelepiped can be computed by multiplying the area of the base by the height.

Let $\vec{v}_1, \vec{v}_2$ and $\vec{v}_3$ be vectors in $\mathbb{R}^3$. For simplicity, let us consider the case when the parallelogram $ABCD$ induced by the vectors $\vec{v}_1$ and $\vec{v}_2$ lies on the xy-plane.



We then have

$$\text{Volume} = \text{Area}(ABCD) \cdot \text{height}$$

From earlier in this lesson, we know the following:

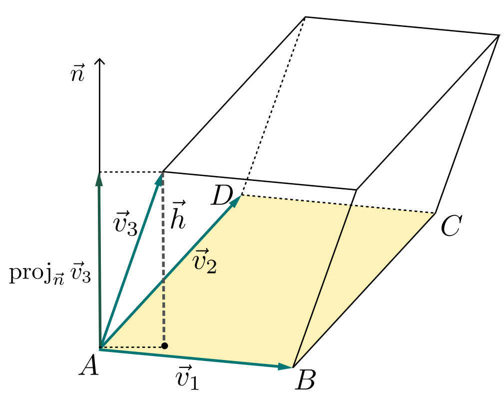
$$\text{Area}(ABCD) = \|\vec{v}_1\| \cdot \|\vec{v}_2\| \cdot |\sin\theta|$$

Also, we have previously proved the following curious relation between the cross product and the area of the parallelogram induced by the two vectors in $\mathbb{R}^3$:

$$\|\vec{v}_1 \times \vec{v}_2\| = \|\vec{v}_1\| \cdot \|\vec{v}_2\| \cdot |\sin\theta|$$

Thus the area of the base is equal to $\|\vec{v}_1 \times \vec{v}_2\|$.

Next, we need to compute the height, $||\vec{h}||$, of our parallelepiped. $||\vec{h}||$ is equal to the projection of vector $\vec{v}_3$ onto the normal vector $\vec{n} = \vec{v}_1 \times \vec{v}_2$.



Thus

$$\|\vec{h}\| = \|\operatorname{proj}_{\vec{n}} \vec{v}_3\| = \frac{|\vec{n} \cdot \vec{v}_3|}{\|\vec{n}\|} = \frac{|(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3|}{\|\vec{v}_1 \times \vec{v}_2\|}$$

Now we are able to compute the volume V as follows:

$$\begin{aligned} V &= \text{Area}(ABCD) \times \|\vec{h}\| \\ &= \|\vec{v}_1 \times \vec{v}_2\| \cdot \frac{|(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3|}{\|\vec{v}_1 \times \vec{v}_2\|} \\ &= |(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3| \end{aligned}$$

□

Making Connections

Observe how the volume of the parallelepiped induced by $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$ changes with respect to other related values as you drag the points P, Q, and R to different locations.

External resource: https://www.geogebra.org/material/iframe/id/qe5xfpfj/

Created with `https://www.geogebra.org`.

Exercise 2

A question appears in Mobius

The n-Volume of a Parallelotope

Formula for the n-Volume

In this section, we will generalize volumes beyond $\mathbb{R}^3$ by studying n-volumes in $\mathbb{R}^n$ for $n \geq 4$.

The ideas that we developed about the determinant and volume generalize as follows:

Proposition 3
Let

$$\vec{v}_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix}, \dots, \vec{v}_n = \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{bmatrix}$$

be vectors in $\mathbb{R}^n$. Then the n-volume V of the n-dimensional parallelotope induced by $\vec{v}_1, \cdots, \vec{v}_n$ is equal to

$$V = \left|\det \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix}\right| = \left|\det \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix}\right|$$

Example

Compute the 4-volume of the parallelotope in $\mathbb{R}^4$ induced by the vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 0 \end{bmatrix}$$

Solution

The 4-volume V of the parallelotope can be computed as follows:

$$\begin{aligned} V &= |\det \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 \end{bmatrix}| \\ &= \left|\det \begin{bmatrix} 1 & -1 & 0 & 2 \\ -1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 3 \\ 0 & 2 & -1 & 0 \end{bmatrix}\right| \end{aligned}$$

In order to compute the determinant, we will apply a cofactor expansion. To simplify our calculations, we will pick the row or column with the most zeros. In this case, we see that the 4th row and the 3rd column both contain two zeros. We pick the 4th row:

$$\begin{aligned} \det \begin{bmatrix} 1 & -1 & 0 & 2 \\ -1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 3 \\ 0 & 2 & -1 & 0 \end{bmatrix} &= -0 \cdot \det \begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 0 & 3 \end{bmatrix} + 2 \cdot \det \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 1 \\ 1 & 0 & 3 \end{bmatrix} \\ &\quad - (-1) \cdot \det \begin{bmatrix} 1 & -1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix} + 0 \cdot \det \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\ &= 2 \cdot \det \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 1 \\ 1 & 0 & 3 \end{bmatrix} + \det \begin{bmatrix} 1 & -1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix} \end{aligned}$$

We are left with the computation of two 3-by-3 determinants. For the first one, it is natural to use the cofactor expansion across the second column:

$$\det\begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 1 \\ 1 & 0 & 3 \end{bmatrix} = 1\cdot\det\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = 1\cdot(1\cdot 3 - 1\cdot 2) = 1$$

For the second one, we can also use the cofactor expansion across the second column, as it contains a zero in the middle:

$$\begin{aligned}\det\begin{bmatrix} 1 & -1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix} &= -(-1)\cdot\det\begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix} - 1\cdot\det\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \\ &= (-1\cdot 3 - 1\cdot 1) - (1\cdot 1 - (-1)\cdot 2) \\ &= -7\end{aligned}$$

Coming back to the original calculation of the 4-by-4 determinant, we find that

$$\begin{aligned}\det\begin{bmatrix} 1 & -1 & 0 & 2 \\ -1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 3 \\ 0 & 2 & -1 & 0 \end{bmatrix} &= 2\cdot\det\begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 1 \\ 1 & 0 & 3 \end{bmatrix} + \det\begin{bmatrix} 1 & -1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix} \\ &= 2\cdot 1 + (-7) \\ &= -5\end{aligned}$$

We conclude that the 4-volume of our parallelotope is $V = |-5| = 5$.

Exercise

A question appears in Mobius

Geometric Interpretation of the Determinant

So far, we have seen that the determinant can be used to determine the area of a parallelogram, volume of a parallelepiped, and n-volume of a parallelotope. In this section, we will further explore the information that is contained in the value of the determinant.

Recall that we can associate an $n \times n$ matrix $[L]$ to a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^n$. These linear mappings are often used to represent geometrical transformations in $\mathbb{R}^n$ such as rotations, reflections, and shears. Finding $\det[L]$ gives us clues about the behaviour of the mapping L.

Magnitude

We have seen that the magnitude of a determinant can be interpreted as an n-volume. Given a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^n$, $|\det[L]|$ has a similar interpretation: as the following proposition suggests, we can view $|\det[L]|$ as the n-volume scaling factor of the mapping L.

Proposition 4
Let $\vec{v}_1, \ldots, \vec{v}_n$ be vectors in $\mathbb{R}^n$ and let V denote the n-volume of the parallelotope induced by these vectors. Let $L\colon \mathbb{R}^n \to \mathbb{R}^n$ be a linear mapping. Then the n-volume of a parallelotope induced by the vectors $L(\vec{v}_1), \ldots, L(\vec{v}_n)$ is equal to $|\det[L]| \cdot V$.

Example 1

Let's see an example of how the area of the parallelogram induced by

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

changes as we apply a linear mapping L to it.

Consider the linear mapping $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $L(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$. Then its standard matrix is

$$[L] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

The determinant of this matrix is equal to 2, which means that L scales areas by a factor of 2.

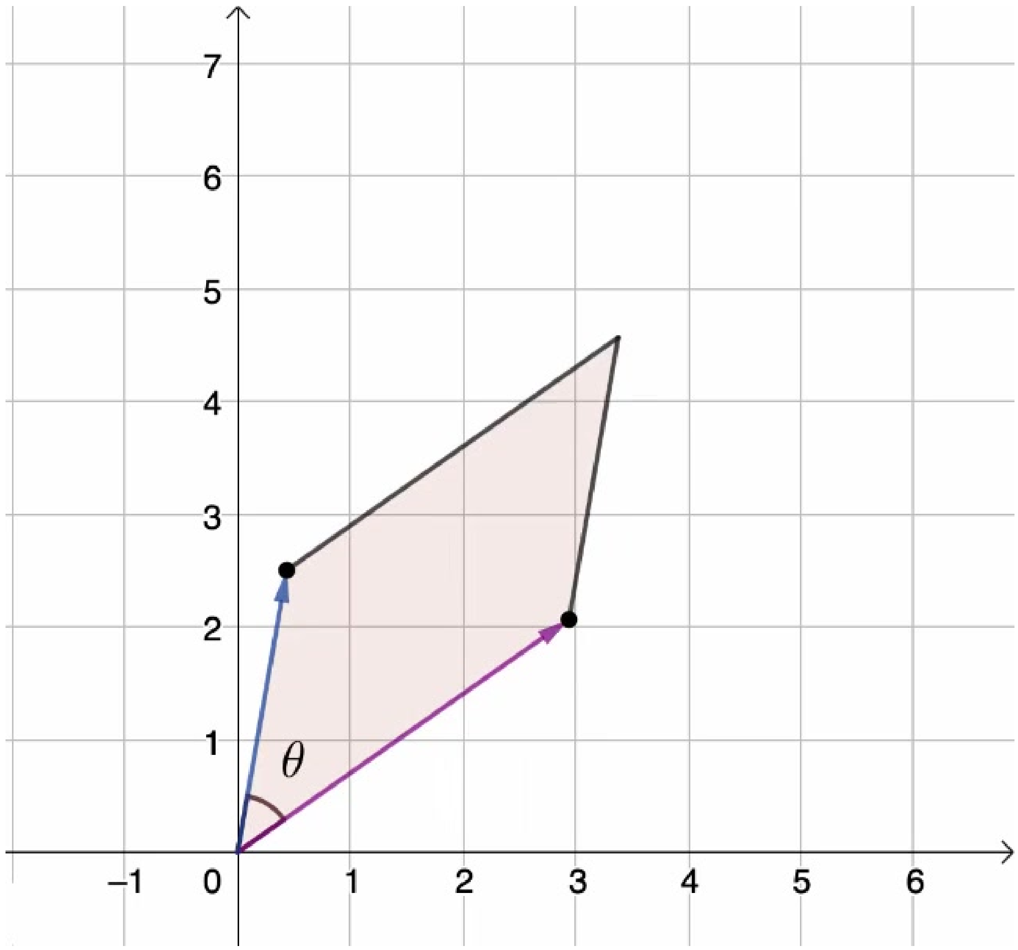
We have

$$L(\vec{v}_1) = \begin{bmatrix} -1 \\ 3 \end{bmatrix} \quad \text{and} \quad L(\vec{v}_2) = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

The parallelogram induced by these two vectors has area 10, which is twice the area of the original parallelogram. Indeed, the action of L changes the parallelogram of area 5 into a parallelogram of area $2 \cdot 5 = 10$.

The following animation illustrates this transformation.

A video appears here.



Vanishing

What does it mean for an n-volume to be zero? Previously in this lesson we saw that:

- the area of a parallelogram is equal to zero whenever the two vectors that induce it lie on the same line passing through the origin, and
- the volume of a parallelepiped is equal to zero whenever the three vectors that induce it lie on the same plane (this is also true for vectors that lie on the same line).

Notice that, in both cases, the vectors are linearly dependent. It turns out that this is a natural criterion for determining whether the n-volume of a parallelotope is equal to zero.

Proposition 5
Let $\vec{v}_1, \ldots, \vec{v}_n$ be vectors in $\mathbb{R}^n$. Then the n-volume of the parallelotope induced by the vectors $\vec{v}_1, \ldots, \vec{v}_n$ is equal to zero if and only if the set $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is linearly dependent.

In the case of a linear mapping L with matrix $[L]$, we can also provide a natural interpretation for the case when $\det[L] = 0$.

Proposition 6
Let $L\colon \mathbb{R}^n \to \mathbb{R}^n$ be a linear mapping. Then $\det[L] \neq 0$ if and only if the mapping L is injective.

Remark
In fact, a much stronger phenomenon holds: since the domain and the codomain of $L\colon \mathbb{R}^n \to \mathbb{R}^n$ are the same, the injectivity of L implies surjectivity of L and vice versa. Maps that are both injective and surjective are called **bijective**.

Example 2

Let's see an example of how the area of the parallelogram induced by

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

changes as we apply a linear mapping L to it.

Consider the linear mapping $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $L(x_1, x_2) = (2x_1 + 2x_2, x_1 + x_2)$. Then its standard matrix is

$$[L] = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$$

The determinant of this matrix is equal to zero, which means that L is not injective.

Notice that, unlike $\vec{v}_1$ and $\vec{v}_2$,

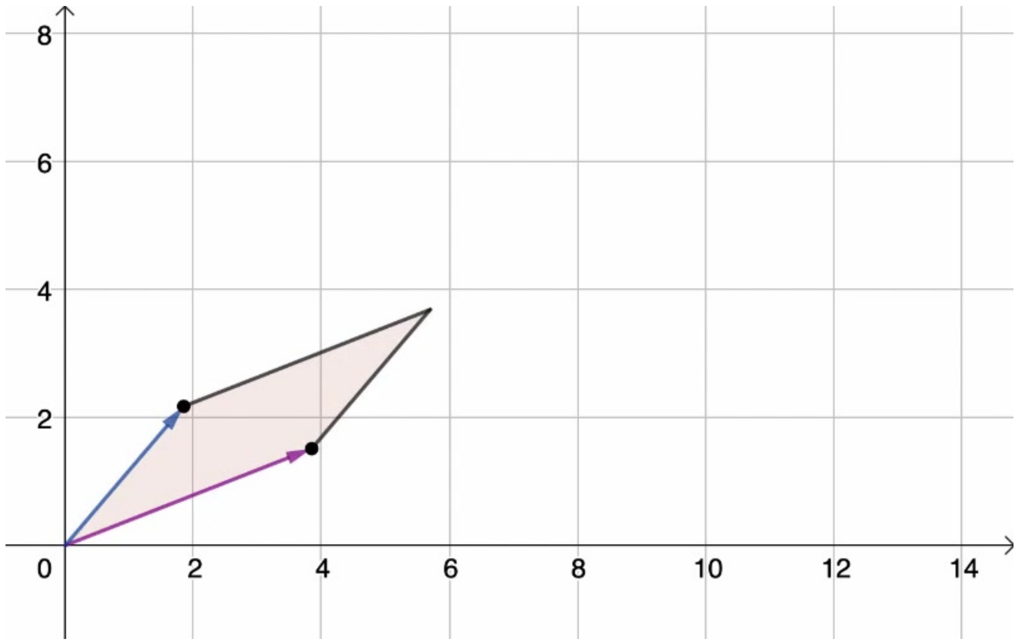
$$L(\vec{v}_1) = \begin{bmatrix} 6 \\ 3 \end{bmatrix} \quad \text{and} \quad L(\vec{v}_2) = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

are on the same line, so the area of the new parallelogram is equal to zero. This observation agrees with the fact that the absolute value of $\det[L]$ is the scaling factor of L.

Indeed, the action of L changes the parallelogram of area 5 into a parallelogram of area $0 \cdot 5 = 0$.

The following animation illustrates this transformation.

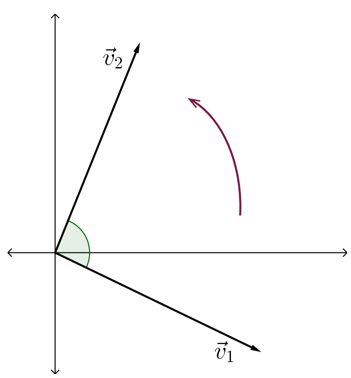
A video appears here.



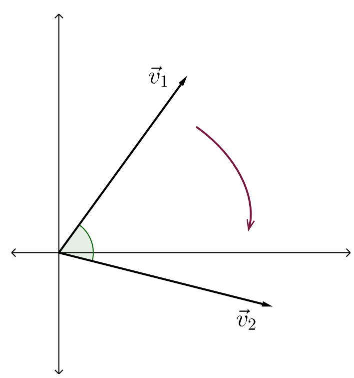
Sign

When we are calculating n-volumes, we always take the absolute value of the determinant; however, the sign of the determinant also provides us with useful information about the mapping.

For simplicity, let's consider non-zero vectors $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^2$. The direction of rotation to get from $\vec{v}_1$ to $\vec{v}_2$ in the shortest way possible can be either counterclockwise or clockwise.



Counterclockwise rotation

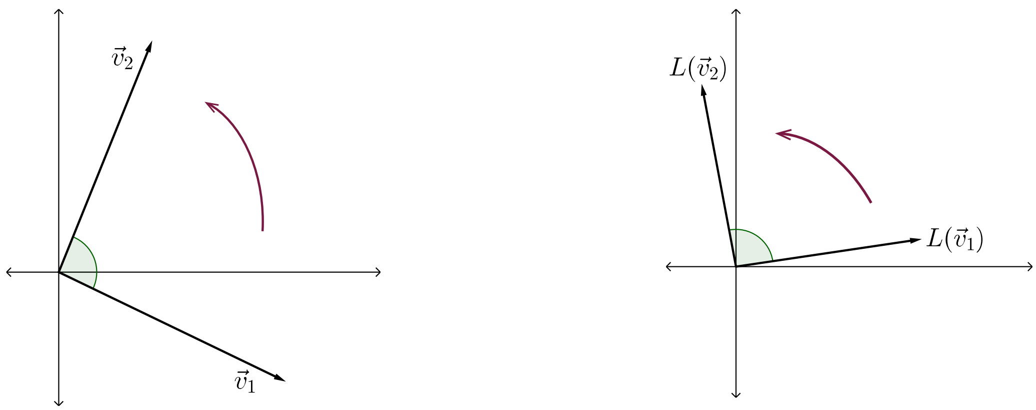


Clockwise rotation

It turns out that the determinant $\det[\vec{v}_1\ \vec{v}_2]$ is positive in the first case and negative in the second case.

If we shift our view towards linear mappings, then a linear mapping L maps $\vec{v}_1$ to $L(\vec{v}_1)$ and $\vec{v}_2$ to $L(\vec{v}_2)$. There are two possibilities.

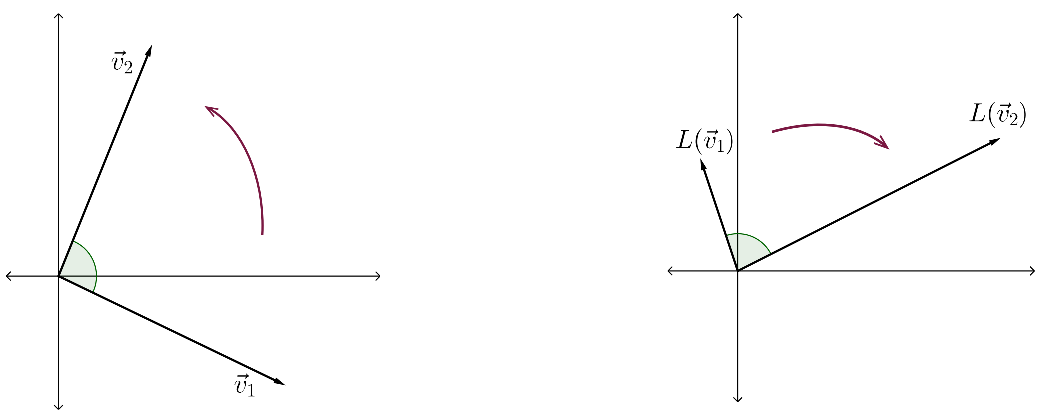
First, the direction of rotation to get from $L(\vec{v}_1)$ to $L(\vec{v}_2)$ in the shortest way possible can be the same as the direction of rotation to get from $\vec{v}_1$ to $\vec{v}_2$ in the shortest way possible:



$\det[L] > 0$

This occurs when $\det[L]$ is positive, and we say that the mapping L **preserves orientation**.

Second, the direction of rotation to get from $L(\vec{v}_1)$ to $L(\vec{v}_2)$ in the shortest way possible can be opposite to the direction of rotation for getting from $\vec{v}_1$ to $\vec{v}_2$ in the shortest way possible:



$\det[L] < 0$

This occurs when $\det[L]$ is negative, and we say that the mapping L **does not preserve orientation**.

We can generalize the idea of orientation to $\mathbb{R}^n$ to get the following result:

Proposition 7
Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a linear mapping with associated matrix $[L]$.

1. If $\det[L] > 0$, then L preserves orientation.
2. If $\det[L] < 0$, then L does not preserve orientation.
3. If $\det[L] = 0$, then no information about orientation is gained.

Example 3

Let's see an example of how the area of the parallelogram induced by

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

changes as we apply a linear mapping L to it.

Consider the linear mapping $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $L(x_1, x_2) = (x_1 + x_2, -x_2)$. Then its standard matrix is

$$[L] = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$$

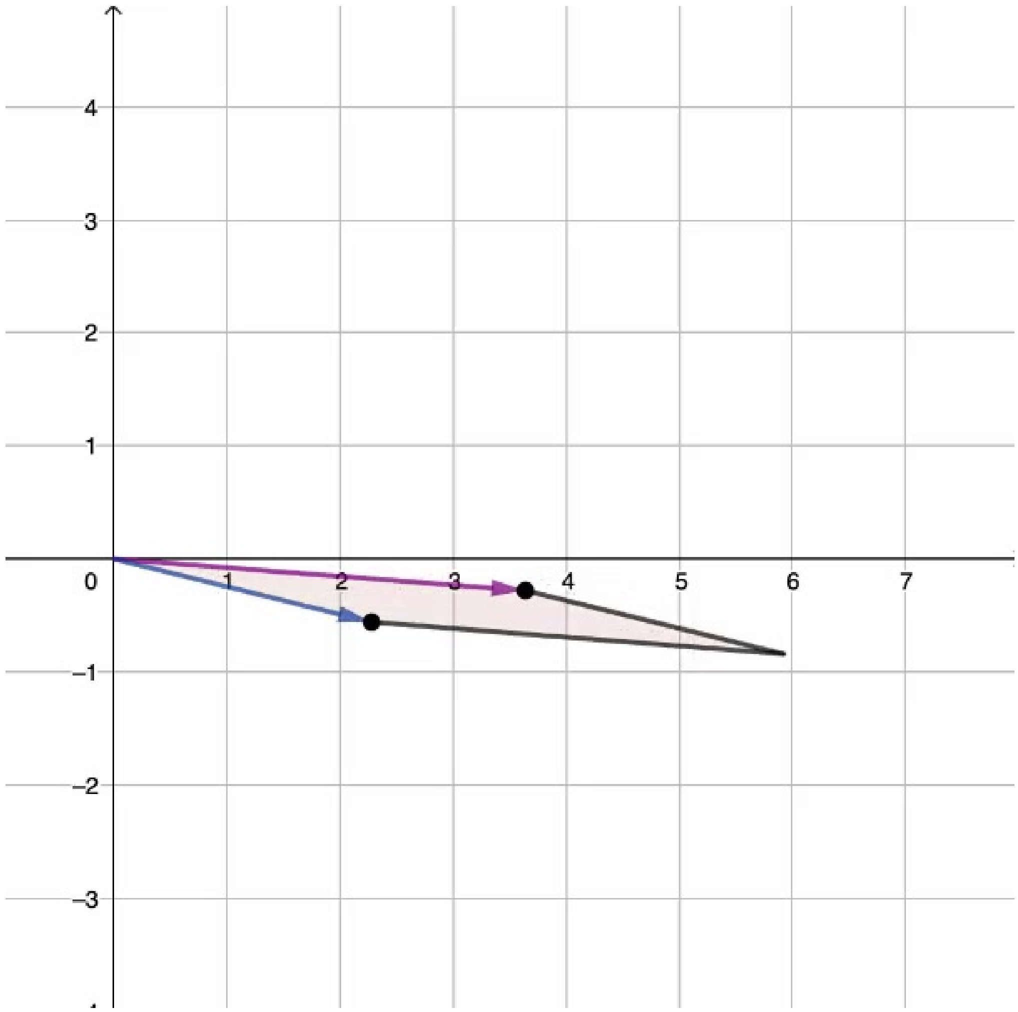
The determinant of $[L]$ is -1, so we expect that $[L]$ does not preserve orientation.

We have

$$L(\vec{v}_1) = \begin{bmatrix} 3 \\ -2 \end{bmatrix} \quad \text{and} \quad L(\vec{v}_2) = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

The following animation illustrates this transformation.

A video appears here.



Notice how the orientation of the starting vectors $\vec{v}_1$ and $\vec{v}_2$ changes as the linear mapping is applied: in the starting configuration, the shortest way to get from $\vec{v}_1$ to $\vec{v}_2$ was rotating clockwise; however, the shortest way to get from $L(\vec{v}_1)$ to $L(\vec{v}_2)$ is counterclockwise. Also notice that the areas of both parallelograms are equal since the scaling factor is 1.

7.5 - Cramers Rule

Cramers Rule

We have seen several useful applications of determinants so far, including checking whether a matrix is invertible and calculating volumes. Here, we will see that determinants can also be used to find a solution to a system of linear equations.

Consider the system of linear equations $A\vec{x} = \vec{b}$ where A is an $n \times n$ matrix. If A is invertible, the system has unique solution $\vec{x} = A^{-1}\vec{b}$.

At the very end of the lesson on finding determinants via cofactor expansion, we saw a theorem which states that, for every invertible matrix A, the equality $A^{-1} = \dfrac{1}{\det(A)}\operatorname{adj}(A)$ holds.

Substituting $\frac{1}{\det(A)}\operatorname{adj}(A)$ for A^{-1} into the expression $\vec{x} = A^{-1}\vec{b}$, we find

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{\det(A)}\operatorname{adj}(A)\vec{b} = \frac{1}{\det(A)}\begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

where C_{ij} is the (i, j)-cofactor of A.

From this, we see that each entry in the column vector is given by

$$x_i = \frac{b_1C_{1i} + \cdots + b_nC_{ni}}{\det(A)}$$

Let's examine the numerator more closely. Recall that the determinant of an $n \times n$ matrix A is given by

$$\det(A) = a_{1i}C_{1i} + a_{2i}C_{2i} + \cdots + a_{ni}C_{ni}$$

where i is a fixed column of A.

If we set $a_{1i} = b_1, \ldots, a_{ni} = b_n$, we find $b_1C_{1i} + \cdots + b_nC_{ni}$: this means that $b_1C_{1i} + \cdots + b_nC_{ni}$ is the determinant of the matrix A whose i-th column has been replaced by $\vec{b}$. Let's call this matrix A_i. It will be of the form

$$A_i = \begin{bmatrix} a_{11} & \cdots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n(i-1)} & b_n & a_{n(i+1)} & \cdots & a_{nn} \end{bmatrix}$$

We can thus write $x_i = \frac{b_1C_{1i} + \cdots + b_nC_{ni}}{\det(A)} = \frac{\det(A_i)}{\det(A)}$.

We can thus determine the solution $\vec{x}$ of the system $A\vec{x} = \vec{b}$ using this approach.

Let's state it formally.

Theorem 1: Cramer's Rule
Let A be an $n \times n$ invertible matrix. Then, the unique solution to $A\vec{x} = \vec{b}$ is given by

$$x_i = \frac{\det(A_i)}{\det(A)} \quad \text{for } 1 \leq i \leq n$$

where A_i is the matrix obtained from A by replacing the i-th column by $\vec{b}$.

How to Use Cramer's Rule

Example

Use Cramer's Rule to solve the system of linear equations

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

Solution

The coefficient matrix of this system is $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix}$. In order to use Cramer's Rule, we must first check whether A is invertible.

Using cofactor expansion along the second row, we find

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \end{vmatrix} \\ &= (-1)\begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} + (-1)\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} \\ &= (-1)(-2-3) + (-1)(1-2) \\ &= 6 \\ &\neq 0 \end{aligned}$$

so A is invertible and Cramer's Rule applies.

From our system of linear equations $A\vec{x} = \vec{b}$, we observe that $\vec{b} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$.

Create matrices A_1, A_2, and A_3 from our original matrix A, by replacing the respective column of A by $\vec{b}$:

- $A_1 = \begin{bmatrix} 2 & 2 & 3 \\ 3 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix}$
- $A_2 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 1 & -1 \end{bmatrix}$
- $A_3 = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 0 & 3 \\ 1 & 1 & 1 \end{bmatrix}$

The matrices A_1, A_2, and A_3 have the following determinants:

- $\det(A_1) = 15$
- $\det(A_2) = -6$
- $\det(A_3) = 3$

By Cramer's Rule, the unique solution to the system of linear equations is given by

- $x_1 = \dfrac{\det(A_1)}{\det(A)} = \dfrac{15}{6} = \dfrac{5}{2}$
- $x_2 = \dfrac{\det(A_2)}{\det(A)} = \dfrac{-6}{6} = -1$
- $x_3 = \dfrac{\det(A_3)}{\det(A)} = \dfrac{3}{6} = \dfrac{1}{2}$

Exercise

A question appears in Mobius

Unit 8

Diagonalization

8.1 - Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

Intuition for Eigenvalues and Eigenvectors

In the unit on linear mappings, we discussed the geometric interpretation of these mappings and saw some common transformations such as rotations, reflections, and stretching/compression.

Given a mapping with matrix A, we are often interested in vectors which are affected by the mapping in a particular way: vectors which are sent to **scalar multiples of themselves** by A. Let's see how this looks geometrically.

Making Connections

1. Select the checkbox to observe how a stretch/compression mapping with matrix A transforms:
   - vector $\vec{u}$,
   - vector $\vec{z}$, and
   - the area of the rectangle formed from vectors $\vec{v}$ and $\vec{w}$.
2. Now deselect the checkbox, and try changing the positions of $\vec{v}$ and $\vec{w}$. Select the checkbox again to see the results on your new rectangle.

External resource: https://www.geogebra.org/material/iframe/id/kqt9qvkt/

You may have noticed that:

- the vector $\vec{u}$ is sent by A to a scalar multiple of itself: in this case, to the vector $\frac{1}{2}\vec{u}$, and
- the vector $\vec{z}$ is **not** sent by A to a scalar multiple of itself.
- the area of the rectangle formed by vectors $\vec{v}$ and $\vec{w}$ does not change.

We say that the vector $\vec{u}$ is an **eigenvector** of A with **eigenvalue** $\lambda = \dfrac{1}{2}$ and that the vector $\vec{z}$ is not an eigenvector of A.

Our goal for the rest of this unit will be to find the eigenvalues and eigenvectors for a given matrix.

Definition of Eigenvalues and Eigenvectors

Now that we have some intuition for eigenvalues and eigenvectors, let's define them more formally.

For $A \in M_{n\times n}(\mathbb{R})$, we are interested in when $A\vec{x} = \lambda\vec{x}$ for a non-zero vector $\vec{x}$ and scalar $\lambda \in \mathbb{R}$. We require that $\vec{x} \neq \vec{0}$ as otherwise the equation becomes $A\vec{0} = \lambda\vec{0}$, which is true for every scalar λ, and isn't very interesting.

> Definitions
> For $A \in M_{n\times n}(\mathbb{R})$, a scalar λ is an **eigenvalue** of A if $A\vec{x} = \lambda\vec{x}$ for some non-zero vector $\vec{x} \in \mathbb{R}^n$. The vector $\vec{x}$ is called an **eigenvector** of A corresponding to λ.

Example 1

Let $A = \begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix}$ and $\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

1. Show that $\vec{x}$ is an eigenvector of A.
2. Find its associated eigenvalue, λ.

Solution - Part A

We have

$$A\vec{x} = \begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1\vec{x},$$

so $\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of A.

Solution - Part B

As we can see in the above work, the corresponding eigenvalue is $\lambda = 1$.

Example 2

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a reflection in the x_2-axis.

1. Find A, the standard matrix of T.
2. Find the eigenvalues and eigenvectors of A.

Solution - Part A

We know that T is a linear transformation, so $A = [T] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ is the standard matrix of T.

Solution - Part B

Thinking geometrically, we see that the reflection of $\vec{e}_1$ in the x_2-axis is $-\vec{e}_1$, that is,

$$A\vec{e}_1 = -\vec{e}_1$$

so $\lambda = -1$ is an eigenvalue of A with corresponding eigenvector $\vec{e}_1$.

Similarly,

$$A\vec{e}_2 = \vec{e}_2$$

so $\lambda = 1$ is an eigenvalue of A with corresponding eigenvector $\vec{e}_2$.

It turns out that $\lambda = -1$ and $\lambda = 1$ are the only eigenvalues of A. Later in this unit we will understand why this is the case.

Remark
In general, finding eigenvalues and eigenvectors cannot be done by inspection. We will learn a systematic way of doing this soon.

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

We end this lesson with a final observation: if λ is an eigenvalue of A with corresponding eigenvector $\vec{x} \neq \vec{0}$, then $A\vec{x} = \lambda\vec{x}$. For any scalar $t \neq 0$, we have that

$$t\vec{x} \neq \vec{0} \quad \text{and} \quad A(t\vec{x}) = tA(\vec{x}) = t(\lambda\vec{x}) = \lambda(t\vec{x})$$

So $t\vec{x}$ is also an eigenvector of A corresponding to λ for any non-zero $t \in \mathbb{R}$. In other words, any non-zero scalar multiple of an eigenvector of A is also an eigenvector of A corresponding to the eigenvalue λ.

8.2 - Finding Eigenvalues and Eigenvectors

Finding Eigenvalues and Eigenvectors of a Matrix

Remember, our goal is to have a method for systematically finding the eigenvalues and eigenvectors of a matrix $A \in M_{n\times n}(\mathbb{R})$.

Let's start by observing that λ is an eigenvalue of A with corresponding eigenvector $\vec{x}$, when

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

Equivalently,

$$A\vec{x} - \lambda\vec{x} = \vec{0}$$

Since multiplying by the identity matrix does not affect the equation, we can write this as

$$A\vec{x} - \lambda I\vec{x} = \vec{0}$$

Factoring out $\vec{x}$, we have that λ is an eigenvalue of A with corresponding eigenvector $\vec{x}$ when

$$(A - \lambda I)\vec{x} = \vec{0}$$

We will therefore consider the homogeneous system $(A - \lambda I)\vec{x} = \vec{0}$.

Notice that the matrix $B = A - \lambda I$ cannot be invertible, for otherwise the equation $B\vec{x} = \vec{0}$ would imply that $\vec{x} = B^{-1}\vec{0} = \vec{0}$, in contradiction to our assumption that $\vec{x} \neq \vec{0}$. Since $A - \lambda I$ cannot be invertible, it follows from the theorem on Invertibility and Determinant stated in an earlier lesson that $\det(A - \lambda I) = 0$.

Characteristic Polynomial

The expression $\det(A - \lambda I)$ will appear very frequently, so we give it a special name.

Definition
The **characteristic polynomial**, denoted $C_A(\lambda)$, of $A \in M_{n\times n}(\mathbb{R})$ is

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

and λ is an eigenvalue of A if and only if $C_A(\lambda) = 0$.

As we will see, $C_A(\lambda)$ is a polynomial. Since $A \in M_{n\times n}(\mathbb{R})$, $C_A(\lambda)$ will have real coefficients, but may have non-real roots. In this course, the characteristic polynomial will always have at least one real root.

Example 1

Let $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Find the characteristic polynomial, $C_A(\lambda)$, of A.

Solution

We start by calculating $A - \lambda I$:

$$A - \lambda I = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & 1 \\ 1 & -1-\lambda \end{bmatrix}$$

Then,

$$\begin{aligned} C_A(\lambda) &= \begin{vmatrix} 1-\lambda & 1 \\ 1 & -1-\lambda \end{vmatrix} \\ &= (1-\lambda)(-1-\lambda) - (1)(1) \\ &= \lambda^2 - 2 \end{aligned}$$

Exercise 1

A question appears in Mobius

Example 2

Let $A = \begin{bmatrix} -2 & 2 & 3 \\ -9 & 7 & 5 \\ -5 & 2 & 6 \end{bmatrix}$. Find the characteristic polynomial of A.

Solution

We start by calculating $A - \lambda I$:

$$A - \lambda I = \begin{bmatrix} -2 & 2 & 3 \\ -9 & 7 & 5 \\ -5 & 2 & 6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} -2-\lambda & 2 & 3 \\ -9 & 7-\lambda & 5 \\ -5 & 2 & 6-\lambda \end{bmatrix}$$

Then, using cofactor expansion along the first row, we have

$$\begin{aligned} C_A(\lambda) &= \begin{vmatrix} -2-\lambda & 2 & 3 \\ -9 & 7-\lambda & 5 \\ -5 & 2 & 6-\lambda \end{vmatrix} \\ &= (-2-\lambda)\begin{vmatrix} 7-\lambda & 5 \\ 2 & 6-\lambda \end{vmatrix} - 2\begin{vmatrix} -9 & 5 \\ -5 & 6-\lambda \end{vmatrix} + 3\begin{vmatrix} -9 & 7-\lambda \\ -5 & 2 \end{vmatrix} \\ &= (-2-\lambda)[(7-\lambda)(6-\lambda)-(2)(5)] - 2[(-9)(6-\lambda)-(5)(-5)] + 3[(-9)(2)-(7-\lambda)(-5)] \\ &= (-2-\lambda)[\lambda^2 - 13\lambda + 32] - 2[9\lambda - 29] + 3[-5\lambda + 17] \\ &= -\lambda^3 + 11\lambda^2 - 39\lambda + 45 \\ &= -(\lambda-3)^2(\lambda-5) \end{aligned}$$

Remark
Observe that in the last example, the matrix A is 3×3 and the characteristic polynomial, $C_A(\lambda)$, was also of degree 3. This is true in general: for $A \in M_{n\times n}(\mathbb{R})$, $C_A(\lambda)$ will be of degree n.

Exercise 2

A question appears in Mobius

Definition
Let $A \in M_{n\times n}(\mathbb{R})$ with eigenvalue λ_i. The **algebraic multiplicity** of λ_i, denoted a_{λ_i}, is the number of times that λ_i appears as a root of $C_A(\lambda)$.

Example 3

Suppose that an 18×18 matrix A has the characteristic polynomial

$$C_A(\lambda) = (\lambda+7)^2\lambda^5(\lambda-11)^9(\lambda^2+\lambda+1)$$

Find the algebraic multiplicities of the real eigenvalues of A.

Solution

Since $\lambda^2+\lambda+1$ has no real roots, the only real eigenvalues of A are $\lambda_1 = -7$, $\lambda_2 = 0$ and $\lambda_3 = 11$. We see that

- $\lambda_1 = -7$ appears as a root 2 times, so $a_{\lambda_1} = 2$;
- $\lambda_2 = 0$ appears as a root 5 times, so $a_{\lambda_2} = 5$; and
- $\lambda_3 = 11$ appears as a root 9 times, so $a_{\lambda_3} = 9$.

Exercise 3

A question appears in Mobius

Algorithm for Finding Eigenvalues and Eigenvectors

Now that we are able to calculate the characteristic polynomial of a matrix A, we can formalize our procedure for finding the eigenvalues and corresponding eigenvectors of A.

Algorithm: Finding Eigenvalues and Eigenvectors
Let A be an $n \times n$ matrix.

1. Find the roots $\lambda_1, \ldots, \lambda_k$ of $C_A(\lambda) = \det(A - \lambda I)$by setting $C_A(\lambda) = 0$.
2. For each eigenvalue λ_i, find the solution set corresponding to the homogeneous system $(A - \lambda_i I)\vec{x} = \vec{0}$. Then the eigenvectors of λ_i are non-zero vectors that belong to the solution set.

Example 4

Let $A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$.

1. Find the eigenvalues of A.
2. Find their corresponding eigenvectors.

Solution - Part A

We start by finding the characteristic polynomial of A:

$$\begin{aligned} C_A(\lambda) &= \det(A - \lambda I) \\ &= \begin{vmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{vmatrix} \\ &= (1-\lambda)(4-\lambda) - 2(-1) \\ &= 4 - 5\lambda + \lambda^2 + 2 \\ &= \lambda^2 - 5\lambda + 6 \\ &= (\lambda - 2)(\lambda - 3) \end{aligned}$$

Next, we find the roots of $C_A(\lambda)$:

$$\begin{aligned} C_A(\lambda) = 0 &\Leftrightarrow (\lambda - 2)(\lambda - 3) = 0 \\ &\Leftrightarrow \lambda = 2 \text{ or } \lambda = 3 \end{aligned}$$

Thus, the roots of $C_A(\lambda)$ are $\lambda_1 = 2$ and $\lambda_2 = 3$, which are the eigenvalues of A.

Solution - Part B

Next, we substitute each eigenvalue into the homogeneous system $(A - \lambda_i I)\vec{x} = \vec{0}$ and solve for $\vec{x}$.

For $\lambda_1 = 2$, we solve $(A - 2I)\vec{x} = \vec{0}$:

$$\begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} \underset{R_2 - R_1}{\sim} \begin{bmatrix} -1 & 2 \\ 0 & 0 \end{bmatrix} \underset{-R_1}{\sim} \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}$$

so

$$\vec{x} = \begin{bmatrix} 2t \\ t \end{bmatrix} = t \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

This gives the eigenvectors of A corresponding to $\lambda_1 = 2$ as

$$t\begin{bmatrix}2\\1\end{bmatrix}, \quad t \neq 0$$

Choosing $t = 1$, we find that $\begin{bmatrix}2\\1\end{bmatrix}$ is an eigenvector associated to the eigenvalue λ_1.

For $\lambda_2 = 3$, we solve $(A - 3I)\vec{x} = \vec{0}$:

$$\begin{bmatrix}-2 & 2\\-1 & 1\end{bmatrix} \underset{R_2 - 1/2R_1}{\sim} \begin{bmatrix}-2 & 2\\0 & 0\end{bmatrix} \underset{-1/2R_1}{\sim} \begin{bmatrix}1 & -1\\0 & 0\end{bmatrix}$$

so

$$\vec{x} = \begin{bmatrix}s\\s\end{bmatrix} = s\begin{bmatrix}1\\1\end{bmatrix}, \quad s \in \mathbb{R}$$

This gives the eigenvectors of A corresponding to $\lambda_2 = 3$ as

$$s\begin{bmatrix}1\\1\end{bmatrix}, \quad s \neq 0$$

Choosing $s = 1$, we find that $\begin{bmatrix}1\\1\end{bmatrix}$ is an eigenvector associated to the eigenvalue λ_2.

This is a good time to check our work:

$$\begin{bmatrix}1 & 2\\-1 & 4\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}4\\2\end{bmatrix} = 2\begin{bmatrix}2\\1\end{bmatrix}$$

and

$$\begin{bmatrix}1 & 2\\-1 & 4\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}3\\3\end{bmatrix} = 3\begin{bmatrix}1\\1\end{bmatrix}$$

Remark
Notice that in the example, we chose a particular value for the free variables t and s. Any non-zero value of t and s will give an eigenvector of A: this means that eigenvectors are **not** unique. It is also important to remember that the zero vector is **not** an eigenvector.

A question appears in Mobius

A question appears in Mobius

Eigenspaces

Finding a Basis for Eigenspaces

Definition
Let λ be an eigenvalue of $A \in M_{n\times n}(\mathbb{R})$. The set containing the zero vector of $\mathbb{R}^n$, together with all eigenvectors of A corresponding to λ is called the **eigenspace** of A corresponding to λ, and is denoted by $E_\lambda(A)$.

It follows from the definition of the eigenspace that $E_\lambda(A) = \mathrm{Null}(A - \lambda I)$. Therefore, $E_\lambda(A)$ is a subspace of $\mathbb{R}^n$.

Remark
Normally, we seek a basis for an eigenspace $E_\lambda(A)$. Once we have a basis, we can construct all of the eigenvectors of A corresponding to λ by taking all non-zero linear combinations of these basis vectors.

From our previous example, the eigenvalues were $\lambda_1 = 2$ and $\lambda_2 = 3$. Hence a basis for $E_{\lambda_1}(A)$ and $E_{\lambda_2}(A)$, respectively, is given by $\left\{\begin{bmatrix}2\\1\end{bmatrix}\right\}$ and $\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$. Note that each eigenspace corresponds to a line in $\mathbb{R}^2$.

A slideshow appears in Mobius.

Slide

Example 1

Let $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Find the eigenvalues of A and, for each one, find a basis for the corresponding eigenspace.

Solution

We start by finding the characteristic polynomial:

$$\begin{aligned}
C_A(\lambda) &= \det(A - \lambda I) \\
&= \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} \underset{\substack{R_1+\lambda R_2 \\ R_3-R_2}}{=} \begin{vmatrix} 0 & 1-\lambda^2 & 1+\lambda \\ 1 & -\lambda & 1 \\ 0 & 1+\lambda & -\lambda-1 \end{vmatrix} \quad \text{by properties of determinant} \\
&= (-1) \begin{vmatrix} (1+\lambda)(1-\lambda) & 1+\lambda \\ 1+\lambda & -(1+\lambda) \end{vmatrix} \quad \text{by cofactor expansion along first column} \\
&= (-1)(1+\lambda)^2 \begin{vmatrix} 1-\lambda & 1 \\ 1 & -1 \end{vmatrix} \\
&= (-1)(\lambda+1)^2((1-\lambda)(-1)-(1)(1)) \\
&= (-1)(\lambda+1)^2(\lambda-2)
\end{aligned}$$

The roots of $C_A(\lambda)$ are $\lambda_1 = -1$ and $\lambda_2 = 2$.
Hence, the eigenvalues of A are $\lambda_1 = -1$ and $\lambda_2 = 2$.

Slide

Example 1 Continued

Let $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Find the eigenvalues of A and, for each one, find a basis for the corresponding eigenspace.

Solution

For $\lambda_1 = -1$, we solve the homogeneous system $(A - \lambda_1 I)\vec{x} = \vec{0}$:

$$(A - (-1)I)\vec{x} = (A + I)\vec{x} = \vec{0}$$

$$A + I = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \underset{\substack{R_2 - R_1 \\ R_3 - R_1}}{\sim} \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\text{RREF}}$$

The general solution is

$$\vec{x} = \begin{bmatrix} -s - t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad s, t \in \mathbb{R}$$

Hence a basis for $E_{\lambda_1}(A)$ is

$$B_1 = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Slide

Example 1 Continued

Let $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Find the eigenvalues of A and, for each one, find a basis for the corresponding eigenspace.

Solution

For $\lambda_2 = 2$, we solve the homogeneous system $(A - \lambda_2 I)\vec{x} = \vec{0}$:

$$(A - 2I)\vec{x} = \vec{0}$$

$$A - 2I = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \sim \underbrace{\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}}_{\text{RREF}}$$

The general solution is

$$\vec{x} = \begin{bmatrix} t \\ t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

Hence a basis for $E_{\lambda_2}(A)$ is given by

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Slide

Important Remark

Remark
Eigenvectors are not unique.

Eigenvectors all live inside their eigenspaces, and eigenspaces are subspaces. This means that **any** vector in the eigenspace associated to a given eigenvalue **is also** an eigenvector with that same eigenvalue.

For example, using our given matrix A, the vector $\vec{v} = \begin{bmatrix} -5 \\ 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ satisfies

$$A\vec{v} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} -5 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ -3 \end{bmatrix} = -\vec{v}$$

and is therefore an eigenvector for $\lambda_1 = -1$.

In fact, **any** non-trivial linear combination of vectors in $B_1 = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is an eigenvector with eigenvalue $\lambda_1 = -1$ (the same applies to B_2).
This means that there are **infinitely many** eigenvectors associated to each eigenvalue!

Remark
It's important to note that eigenvectors are not unique: any linear combination of the basis vectors in an eigenspace is also an eigenvector with the same eigenvalue.

In our previous example, $\lambda_1 = -1$ has algebraic multiplicity $a_{\lambda_1} = 2$ and $\lambda_2 = 2$ has algebraic multiplicity $a_{\lambda_2} = 1$.

There is one more natural parameter that we can associate with the eigenvalue λ, namely the dimension of the eigenspace $E_\lambda(A)$ corresponding to λ. We give this parameter a special name.

Definition
Let $A \in M_{n\times n}(\mathbb{R})$ with eigenvalue λ. The **geometric multiplicity** of λ, denoted g_λ, is the dimension of the eigenspace $E_\lambda(A)$.

In the previous example, we have $\dim(E_{\lambda_1}(A)) = 2$ and $\dim(E_{\lambda_2}(A)) = 1$, so the **geometric multiplicities** corresponding toλ_1 and λ_2 are $g_{\lambda_1} = 2$ and $g_{\lambda_2} = 1$, respectively.

The algebraic and geometric multiplicities of an eigenvalue satisfy the following property:

Theorem 1
For any $A \in M_{n\times n}(\mathbb{R})$ and any eigenvalue λ of A, $1 \leq g_\lambda \leq a_\lambda \leq n$.

Example 2

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

1. Find the eigenvalues of A.
2. A basis for each eigenspace.

Solution - Part A

We start by finding the characteristic polynomial:

$$C_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 0 \\ 5 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$$

Thus λ_1 is the only eigenvalue of A and $a_{\lambda_1} = 2$.

Solution - Part B

Next, we solve $(A - I)\vec{x} = \vec{0}$:

$$\begin{bmatrix} 0 & 0 \\ 5 & 0 \end{bmatrix} \underset{1/5R_2}{\sim} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \underset{R_1 \leftrightarrow R_2}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

The general solution is given by

$$\vec{x} = \begin{bmatrix} 0 \\ t \end{bmatrix} = t \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}$$

Therefore

$$\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

is a basis for $E_{\lambda_1}(A)$, and $g_{\lambda_1} = 1 < 2 = a_{\lambda_1}$.

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

8.3 - Diagonalization

Diagonalization

Diagonal Matrices

In this lesson, we will relate our procedure for finding the eigenvalues and eigenvectors of a matrix to the process of diagonalizing this matrix. In order to understand what it means to diagonalize a matrix, we will need a few definitions and results.

Definition
An $n \times n$ matrix

$$\begin{bmatrix} u_1 & 0 & \cdots & 0 \\ 0 & u_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_n \end{bmatrix}$$

is called a **diagonal matrix**. Such a matrix is denoted by $\mathrm{diag}(u_1, u_2, \ldots, u_n)$.

Example 1

The matrices

$$[7], \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

are diagonal. Note that diagonal matrices are both upper and lower triangular matrices.

Diagonal matrices have some particularly nice properties which make them easy to work with:

Theorem 1: Properties of Diagonal Matrices
Let $D = \text{diag}(u_1, \ldots, u_n)$ and $E = \text{diag}(v_1, \ldots, v_n)$ be diagonal matrices. Then

1. $D + E = \text{diag}(u_1 + v_1, \ldots, u_n + v_n)$. That is, the sum of diagonal matrices is the diagonal matrix of the sums of diagonal entries.
2. $DE = \text{diag}(u_1 v_1, \ldots, u_n v_n)$. That is, the product of diagonal matrices is the diagonal matrix of the products of diagonal entries.
3. For any positive integer k, $D^k = \text{diag}(u_1^k, \ldots, u_n^k)$. That is, the positive k-th power of a diagonal matrix is the diagonal matrix with each entry raised to the power of k.
4. If the $u_1, \ldots, u_n$ are all non-zero, then property 3 holds for **every** integer k.

As this theorem tells us, adding, multiplying, and taking powers of diagonal matrices is much simpler than performing those operations on non-diagonal matrices. For this reason, we would like a way to "transform" a given matrix A into a diagonal matrix D which is easier to work with. This process is called **diagonalization**.

Diagonalization

Let's formalize what it means for a matrix to be diagonalizable.

Definition
A matrix $A \in M_{n\times n}(\mathbb{R})$ is **diagonalizable** if there exists an $n \times n$ invertible matrix P and an $n \times n$ diagonal matrix D such that $P^{-1}AP = D$.

If a 2×2 matrix A is diagonalizable, then the matrices P and D can be computed as follows.

For simplicity, assume $A \in M_{2\times 2}(\mathbb{R})$ and that there exists an invertible 2×2 diagonal matrix D such that $P^{-1}AP = D$.

Let $P = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 \end{bmatrix}$ and $D = \text{diag}(\lambda_1, \lambda_2)$.

Since P is invertible, $\{\vec{x}_1, \vec{x}_2\}$ is linearly independent, so $\vec{x}_1 \neq \vec{0}$ and $\vec{x}_2 \neq \vec{0}$.

From $P^{-1}AP = D$ we have $AP = PD$. In other words,

$$A \begin{bmatrix} \vec{x}_1 & \vec{x}_2 \end{bmatrix} = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Multiplying the matrices on both sides, we find that

$$\begin{bmatrix} A\vec{x}_1 & A\vec{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \vec{x}_1 & \lambda_2 \vec{x}_2 \end{bmatrix}$$

Since the two matrices are equal, their columns must be equal, i.e.,

$$A\vec{x}_1 = \lambda_1 \vec{x}_1 \quad \text{and} \quad A\vec{x}_2 = \lambda_2 \vec{x}_2$$

Since $\vec{x}_1 \neq \vec{0}$ and $\vec{x}_2 \neq \vec{0}$, the numbers λ_1 and λ_2 are the eigenvalues of A, and $\vec{x}_1, \vec{x}_2$ are corresponding basis vectors for the eigenspaces of A.

From this derivation, we notice that the matrix D contains the eigenvalues of A along its diagonal and that the matrix P contains the eigenvectors that correspond to the eigenvalues of A as its columns in the same order.

Example 2

Diagonalize the matrix $A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$.

Solution

We start by finding the determinant of the matrix $A - \lambda I$:

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

Thus, the eigenvalues are

- $\lambda_1 = 2$ with algebraic multiplicity $a_{\lambda_1} = 1$, and
- $\lambda_2 = 3$ with algebraic multiplicity $a_{\lambda_2} = 1$.

The bases of corresponding eigenspaces are

- $E_{\lambda_1}(A)$: $\left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$
- $E_{\lambda_2}(A)$: $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

Since bases of each of the above eigenspaces contain only one element, we deduce that their geometric multiplicities are $g_{\lambda_1} = g_{\lambda_2} = 1$. Thus

$$P = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

and

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

That is, an eigenvector of A in the i-th column of P forces the i-th column of D to contain the corresponding eigenvalue. We say that the matrix P **diagonalizes** A to D, i.e., $P^{-1}AP = D$.

We can check our work as follows

$$P^{-1} = \frac{1}{\det P} \operatorname{adj} P = \frac{1}{1} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

and

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -2 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \\ &= D \end{aligned}$$

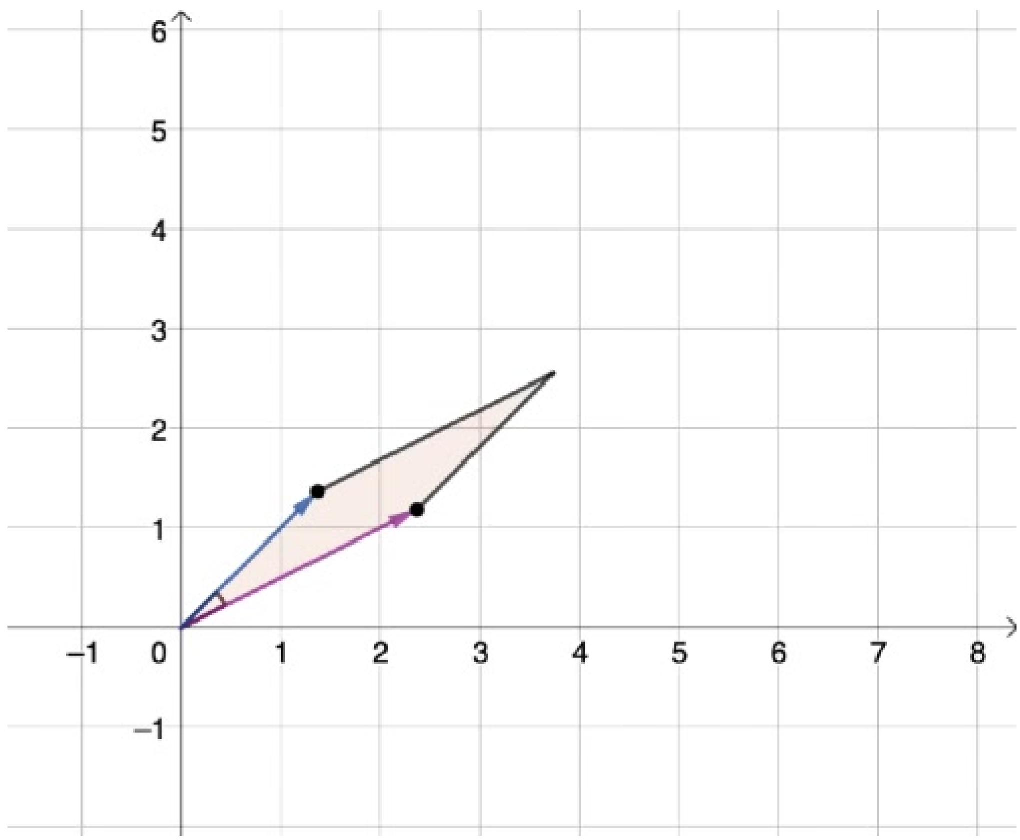
Note that P and D are **not** unique. We could have chosen

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

as well by changing the order in which the eigenvectors appear in P (and therefore in D). Moreover, we can use the vectors in any basis for the eigenspaces of A.

Now, let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the linear mapping whose standard matrix is L. The following video (no audio) demonstrates how the parallelogram induced by the vectors $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ gets transformed under the action of L. More precisely, we see that the parallelogram gets scaled by a factor of 3 in the direction of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and by a factor of 2 in the direction of $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

A video appears here.



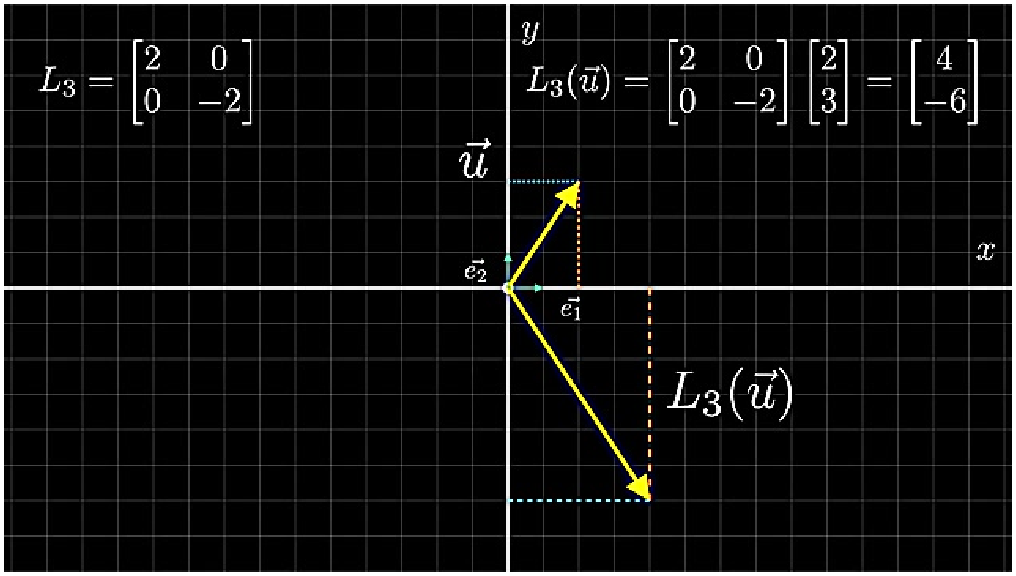
Intuition Behind Diagonalization

Notice that if the standard matrix $[L]$ of $L\colon \mathbb{R}^n \to \mathbb{R}^n$ is diagonal, then it is rather easy to explain how L acts on any vector $\vec{u}$ in $\mathbb{R}^n$. In particular, if $[L] = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, then $L(\vec{u})$ is the result of scaling of $\vec{u}$ by a factor of λ_i in the direction of the i-th standard basis vector $\vec{e}_i$ for all $i = 1, \ldots, n$. In other words,

$$\text{if } \vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \text{then } L(\vec{u}) = \begin{bmatrix} \lambda_1 u_1 \\ \vdots \\ \lambda_n u_n \end{bmatrix}$$

The following video demonstrates this for the case $n = 2$.

A video appears here.

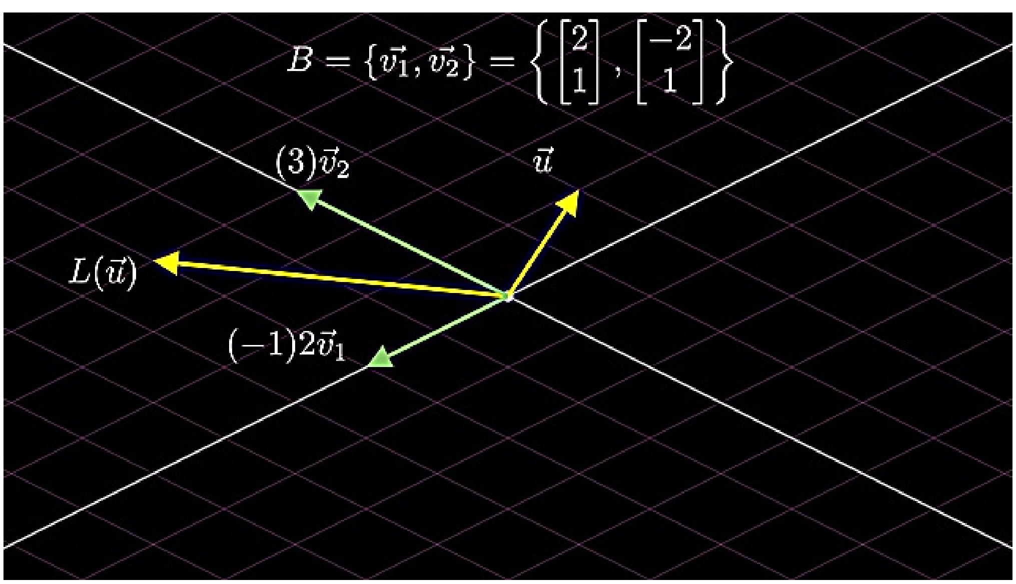


If, however, the standard matrix $[L]$ is not diagonal, then it may not be obvious how L acts on vectors $\vec{u}$ in $\mathbb{R}^n$. Fortunately, if $[L]$ is diagonalizable, we can perform diagonalization so to understand the action of L. In particular, by finding a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ of eigenvectors of $[L]$, we can conclude that $L(\vec{u})$ is the result of scaling of $\vec{u}$ by a factor of λ_i in the direction of the i-th eigenvector $\vec{v}_i$ for all $i = 1, \ldots, n$. In other words,

$$\text{if } [\vec{u}]_{\mathcal{B}} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \text{ then } [L(\vec{u})]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 u_1 \\ \vdots \\ \lambda_n u_n \end{bmatrix}$$

The following video demonstrates this in the case $n = 2$.

A video appears here.



The next example shows how to find a diagonal matrix D and an invertible matrix P from a given matrix A, such that $P^{-1}AP = D$.

A slideshow appears in Mobius.

Slide

Example 3

Let $A = \begin{bmatrix} 3 & -2 & -2 \\ -1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$. Find a diagonal matrix D and an invertible matrix P such that $P^{-1}AP = D$.

Solution

We start by finding the characteristic polynomial:

$$\begin{aligned} C_A(\lambda) &= \det(A - \lambda I) \\ &= \begin{vmatrix} 3-\lambda & -2 & -2 \\ -1 & -\lambda & 0 \\ 2 & 0 & -\lambda \end{vmatrix} \\ &= (-1)^{3+1} 2 \begin{vmatrix} -2 & -2 \\ -\lambda & 0 \end{vmatrix} + (-1)^{3+3}(-\lambda) \begin{vmatrix} 3-\lambda & -2 \\ -1 & -\lambda \end{vmatrix} && \text{by cofactor expansion along third row} \\ &= 2\left((-2)\cdot 0 - (-\lambda)(-2)\right) - \lambda((3-\lambda)(-\lambda) - (-1)(-2)) \\ &= -\lambda(\lambda-1)(\lambda-2) \end{aligned}$$

Hence the eigenvalues of A are

- $\lambda_1 = 0$ with $a_{\lambda_1} = 1$
- $\lambda_2 = 1$ with $a_{\lambda_2} = 1$
- $\lambda_3 = 2$ with $a_{\lambda_3} = 1$

Slide

Example 3 Continued

Let $A = \begin{bmatrix} 3 & -2 & -2 \\ -1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$. Find a diagonal matrix D and an invertible matrix P such that $P^{-1}AP = D$.

Solution

For $\lambda_1 = 0$, we solve the homogeneous system $(A - \lambda_1 I)\vec{x} = \vec{0}$:

$$(A - (0)I)\vec{x} = A\vec{x} = \vec{0}$$

$$A = \begin{bmatrix} 3 & -2 & -2 \\ -1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix} \underset{\substack{R_1+3R_2 \\ R_3+2R_2}}{\sim} \begin{bmatrix} 0 & -2 & -2 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \underset{\substack{(-1/2)R_1 \\ -R_2}}{\sim} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \underset{R_1 \leftrightarrow R_2}{\sim} \underset{\text{RREF}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}}$$

The general solution is

$$\vec{x} = \begin{bmatrix} 0 \\ -s \\ s \end{bmatrix} = s \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \quad s \in \mathbb{R}$$

Hence a basis for $E_{\lambda_1}(A)$ is

$$\left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$$

Slide

Example 3 Continued

Let $A = \begin{bmatrix} 3 & -2 & -2 \\ -1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$. Find a diagonal matrix D and an invertible matrix P such that $P^{-1}AP = D$.

Solution

Analogously, we can determine bases of the eigenspaces corresponding to the eigenvalues $\lambda_2 = 1$ and $\lambda_3 = 2$.

By solving the homogeneous system $(A - \lambda_2 I)\vec{x} = \vec{0}$, we find that a basis for $E_{\lambda_2}(A)$ is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} \right\}$$

Finally, we solve the homogeneous system $(A - \lambda_3 I)\vec{x} = \vec{0}$ and find that a basis for $E_{\lambda_3}(A)$ is

$$\left\{ \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} \right\}$$

Slide

Example 3 Continued

Let $A = \begin{bmatrix} 3 & -2 & -2 \\ -1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$. Find a diagonal matrix D and an invertible matrix P such that $P^{-1}AP = D$.

Solution

We now have everything we need to find a diagonal matrix D and an invertible matrix P such that $P^{-1}AP = D$.

We know that the eigenvalues of A are $\lambda_1 = 0$, $\lambda_2 = 1$, and $\lambda_3 = 2$.

Now let $\vec{v}_1, \vec{v}_2$ and $\vec{v}_3$ be eigenvectors corresponding to λ_1, λ_2, and λ_3, respectively.

$$D = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3) \qquad P = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$$

so that

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad P = \begin{bmatrix} 0 & -1 & 2 \\ -1 & 1 & -1 \\ 1 & -2 & 2 \end{bmatrix}$$

Then the equality $P^{-1}AP = D$ is satisfied.

Therefore, A is diagonalizable.

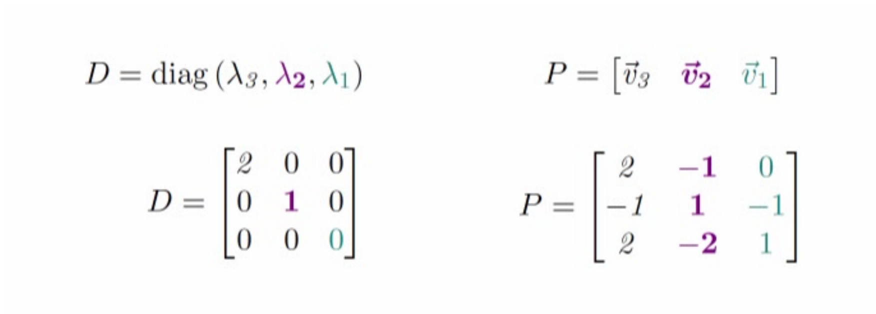
Slide

Important Remark

Remark
The diagonal matrix D and invertible matrix P that satisfy $P^{-1}AP = D$ are **not** unique, because we are allowed to permute diagonal entries of D and respective columns of P .

From our example, we have that the eigenvalues of A are $\lambda_1 = 0$, $\lambda_2 = 1$ and $\lambda_3 = 2$ and $\vec{v}_1, \vec{v}_2$ and $\vec{v}_3$ are eigenvectors corresponding to λ_1, λ_2 and λ_3, respectively.

A video appears here.



Slide

Important Remark (Continued)

In summary, our diagonal matrix D and invertible matrix P are not unique. Below are all of the permutations of diagonal entries of $D = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$, and respective columns of $P = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$, that satisfy $P^{-1}AP = D$.

$D = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$, $P = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$
$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, P = \begin{bmatrix} 0 & -1 & 2 \\ -1 & 1 & -1 \\ 1 & -2 & 2 \end{bmatrix}$$

$D = \operatorname{diag}(\lambda_1, \lambda_3, \lambda_2)$, $P = \begin{bmatrix} \vec{v}_1 & \vec{v}_3 & \vec{v}_2 \end{bmatrix}$
$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, P = \begin{bmatrix} 0 & 2 & -1 \\ -1 & -1 & 1 \\ 1 & 2 & -2 \end{bmatrix}$$

$D = \operatorname{diag}(\lambda_2, \lambda_1, \lambda_3)$, $P = \begin{bmatrix} \vec{v}_2 & \vec{v}_1 & \vec{v}_3 \end{bmatrix}$
$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}, P = \begin{bmatrix} -1 & 0 & 2 \\ 1 & -1 & -1 \\ -2 & 1 & 2 \end{bmatrix}$$

$D = \operatorname{diag}(\lambda_2, \lambda_3, \lambda_1)$, $P = \begin{bmatrix} \vec{v}_2 & \vec{v}_3 & \vec{v}_1 \end{bmatrix}$
$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, P = \begin{bmatrix} -1 & 2 & 0 \\ 1 & -1 & -1 \\ -2 & 2 & 1 \end{bmatrix}$$

$D = \operatorname{diag}(\lambda_3, \lambda_1, \lambda_2)$, $P = \begin{bmatrix} \vec{v}_3 & \vec{v}_1 & \vec{v}_2 \end{bmatrix}$
$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, P = \begin{bmatrix} 2 & 0 & -1 \\ -1 & -1 & 1 \\ 2 & 1 & -2 \end{bmatrix}$$

$D = \operatorname{diag}(\lambda_3, \lambda_2, \lambda_1)$, $P = \begin{bmatrix} \vec{v}_3 & \vec{v}_2 & \vec{v}_1 \end{bmatrix}$
$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, P = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & -1 \\ 2 & -2 & 1 \end{bmatrix}$$

Here is another example of diagonalization.

Example 4

Diagonalize $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

Solution

We start by finding $C_A(\lambda)$; i.e., the determinant of the matrix $A - \lambda I$:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} \\ &= -(\lambda + 1)^2(\lambda - 2) \end{aligned}$$

Thus, the eigenvalues of A are

- $\lambda_1 = -1$ with algebraic multiplicity $a_{\lambda_1} = 2$ and
- $\lambda_2 = 2$ with algebraic mutliplicity $a_{\lambda_2} = 1$.

The bases for the corresponding eigenspaces are

- $E_{\lambda_1}(A)$: $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$
- $E_{\lambda_2}(A)$: $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

Since a basis of $E_{\lambda_1}(A)$ contains two elements, we find that its geometric multiplicity is $g_{\lambda_1} = 2$. For $E_{\lambda_2}(A)$ we see that the geometric multiplicity is $g_{\lambda_2} = 1$.

We take

$$P = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

and

$$D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

We conclude that P diagonalizes A to D, i.e., $P^{-1}AP = D$.

Exercise 1

A question appears in Mobius

Exercise 2

A question appears in Mobius

Determining When a Matrix is Diagonalizable

So far, we have discussed how to diagonalize a matrix A; however, not every square matrix is diagonalizable. Here, we will give a criterion for determining whether a matrix is diagonalizable.

Let's start with an example.

Example 5

Determine whether the matrix

$$A = \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix}$$

is diagonalizable.

Solution

Recall that for

$$A = \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix}$$

the only eigenvalue of A is $\lambda_1 = 1$ with algebraic multiplicity $a_{\lambda_1} = 2$. However, a basis for the eigenspace $E_{\lambda_1}(A)$ is

$$\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

so $g_{\lambda_1} = 1 \neq 2 = a_{\lambda_1}$. Hence A is not diagonalizable as we cannot find two linearly independent eigenvectors of A to form an invertible 2×2 matrix P.

We therefore formulate the following criterion.

Theorem 2: Criterion for Diagonalizability
A matrix $A \in M_{n\times n}(\mathbb{R})$ is diagonalizable if and only if $a_\lambda = g_\lambda$ for each eigenvalue λ of A.

Also, since $1 \leq g_\lambda \leq a_\lambda$, we see that each matrix A that has n distinct eigenvalues must satisfy $a_\lambda = g_\lambda$, because $1 \leq g_\lambda \leq a_\lambda = 1$. It follows from the above theorem that any such matrix is diagonalizable:

Corollary 3
A matrix $A \in M_{n\times n}(\mathbb{R})$ with n distinct eigenvalues is diagonalizable.

Exercise 3

A question appears in Mobius

A question appears in Mobius

A question appears in Mobius

8.4 - Applications

Application to Matrix Powers

Let's see how diagonalization can be applied to simplify taking powers of matrices. Suppose that $A \in M_{n\times n}(\mathbb{R})$ is diagonalizable. Then, as we've seen previously, $P^{-1}AP = D$ for some $n \times n$ invertible matrix P and $n \times n$ diagonal matrix D.

Rearranging this equation gives

$$A = PDP^{-1}$$

Squaring both sides and simplifying gives

$$\begin{aligned} A^2 &= PDP^{-1}PDP^{-1} \\ &= PDIDP^{-1} \\ &= PD^2P^{-1} \end{aligned}$$

Similarly, $A^3 = PD^3P^{-1}$ and more generally, $A^k = PD^kP^{-1}$ for any positive integer k.

Examples and Exercises

Example 1

Find A^k for

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

Solution

We begin with the computation of the characteristic polynomial of A:

$$\begin{aligned} C_A(\lambda) &= \begin{vmatrix} 3-\lambda & -1 \\ 2 & -\lambda \end{vmatrix} \\ &= (3-\lambda)(-\lambda) - (-2) \\ &= \lambda^2 - 3\lambda + 2 \\ &= (\lambda - 1)(\lambda - 2) \end{aligned}$$

Thus $\lambda_1 = 1$ and $\lambda_2 = 2$ are the eigenvalues of A. We see that the algebraic multiplicities are $a_{\lambda_1} = a_{\lambda_2} = 1$. Since A has $n = 2$ distinct eigenvalues, we are guaranteed that A is diagonalizable.

Next, we compute a basis for the eigenspace $E_{\lambda_1}(A)$:

$$A - I = \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix} \underset{R_2 - R_1}{\sim} \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}$$

so

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$

is a basis for $E_{\lambda_1}(A)$.

We proceed with the computation of a basis for the eigenspace $E_{\lambda_2}(A)$:

$$A - 2I = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix} \underset{R_2 - 2R_1}{\sim} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

so

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

is a basis for $E_{\lambda_2}(A)$.

Now, let

$$P = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Thus, we have

$$\begin{aligned} A^k &= PD^kP^{-1} \\ &= \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & 2^k \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2^k \\ 2 & 2^k \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -1 + 2^{k+1} & 1 - 2^k \\ -2 + 2^{k+1} & 2 - 2^k \end{bmatrix} \end{aligned}$$

Note that we can verify our work:

$$A^1 = \begin{bmatrix} -1+2^{1+1} & 1-2^1 \\ -2+2^{1+1} & 2-2^1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix} = A$$

We can also easily compute, for example, A^8:

$$A^8 = \begin{bmatrix} 511 & -255 \\ 510 & -254 \end{bmatrix}$$

Example 2

Find A^k for

$$A = \begin{bmatrix} 3 & -4 \\ -2 & 1 \end{bmatrix}$$

Solution

We begin with the computation of the characteristic polynomial of A:

$$\begin{aligned} C_A(\lambda) &= \begin{vmatrix} 3-\lambda & -4 \\ -2 & 1-\lambda \end{vmatrix} \\ &= (3-\lambda)(1-\lambda) - 8 \\ &= \lambda^2 - 4\lambda + 3 - 8 \\ &= (\lambda-5)(\lambda+1) \end{aligned}$$

Thus $\lambda_1 = -1$ and $\lambda_2 = 5$ are the eigenvalues of A. We see that the algebraic multiplcities are $a_{\lambda_1} = a_{\lambda_2} = 1$. Since A has $n = 2$ distinct eigenvalues, we are guaranteed that A is diagonalizable.

Next, we compute a basis for the eigenspace $E_{\lambda_1}(A)$:

$$A + I = \begin{bmatrix} 4 & -4 \\ -2 & 2 \end{bmatrix} \underset{R_2+1/2R_1}{\sim} \begin{bmatrix} 4 & -4 \\ 0 & 0 \end{bmatrix} \underset{(1/4)R_1}{\sim} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

so

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

is a basis for $E_{\lambda_1}(A)$.

We proceed with the computation of a basis for the eigenspace $E_{\lambda_2}(A)$:

$$A - 5I = \begin{bmatrix} -2 & -4 \\ -2 & -4 \end{bmatrix} \underset{R_2-R_1}{\sim} \begin{bmatrix} -2 & -4 \\ 0 & 0 \end{bmatrix} \underset{(-1/2)R_1}{\sim} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

so

$$\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$$

is a basis for $E_{\lambda_2}(A)$.

Now, let

$$P = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}$$

Then

$$P^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$$

and

$$\begin{aligned} A^k &= PD^kP^{-1} \\ &= \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} (-1)^k & 0 \\ 0 & 5^k \end{bmatrix}\frac{1}{3}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} (-1)^k & (-2)5^k \\ (-1)^k & 5^k \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} (-1)^k + (2)5^k & 2(-1)^k - (2)5^k \\ (-1)^k - 5^k & 2(-1)^k + 5^k \end{bmatrix} \end{aligned}$$

A question appears in Mobius

A question appears in Mobius

Application to Determinant of a Matrix

We can use the eigenvalues of A to find the determinant of A. Suppose the eigenvalues of A are $\lambda_1, \ldots, \lambda_k$ (which are all distinct) with algebraic multiplicities $a_{\lambda_1}, \ldots, a_{\lambda_k}$.

Then

$$a_{\lambda_1} + \cdots + a_{\lambda_k} = n$$

and

$$C_A(\lambda) = \det(A - \lambda I) = (-1)^n(\lambda - \lambda_1)^{a_{\lambda_1}} \cdots (\lambda - \lambda_k)^{a_{\lambda_k}}$$

Taking $\lambda = 0$ gives

$$\begin{aligned}
\det(A) &= (-1)^n(-\lambda_1)^{a_{\lambda_1}} \cdots (-\lambda_k)^{a_{\lambda_k}} \\
&= (-1)^n(-1)^{a_{\lambda_1}+\cdots+a_{\lambda_k}} \lambda_1^{a_{\lambda_1}} \cdots \lambda_k^{a_{\lambda_k}} \\
&= (-1)^n(-1)^n \lambda_1^{a_{\lambda_1}} \cdots \lambda_k^{a_{\lambda_k}} \\
&= \lambda_1^{a_{\lambda_1}} \cdots \lambda_k^{a_{\lambda_k}}
\end{aligned}$$

Thus, $\det(A)$ is the **product** of the eigenvalues of A where each eigenvalue λ of A appears in the product a_λ times.

We also note that A is invertible if and only if 0 is not an eigenvalue of A.

Example and Exercise

Example

Find the determinant of the matrix

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

which has characteristic polynomial $(\lambda + 1)^2(\lambda - 2)$.

Solution

The eigenvalues of A are $\lambda_1 = -1$ (with algebraic multiplicity $a_{\lambda_1} = 2$) and $\lambda_2 = 2$ (with algebraic multiplicity $a_{\lambda_2} = 1$). Thus,

$$\det(A) = \lambda_1^{a_{\lambda_1}} \lambda_2^{a_{\lambda_2}} = (-1)^2 2^1 = 2$$

A question appears in Mobius

A question appears in Mobius