Chapter 1: Systems of linear equations

Linear equations

First example: a linear equation in two variables

Consider the equation

$$2x + 5y = 7.$$

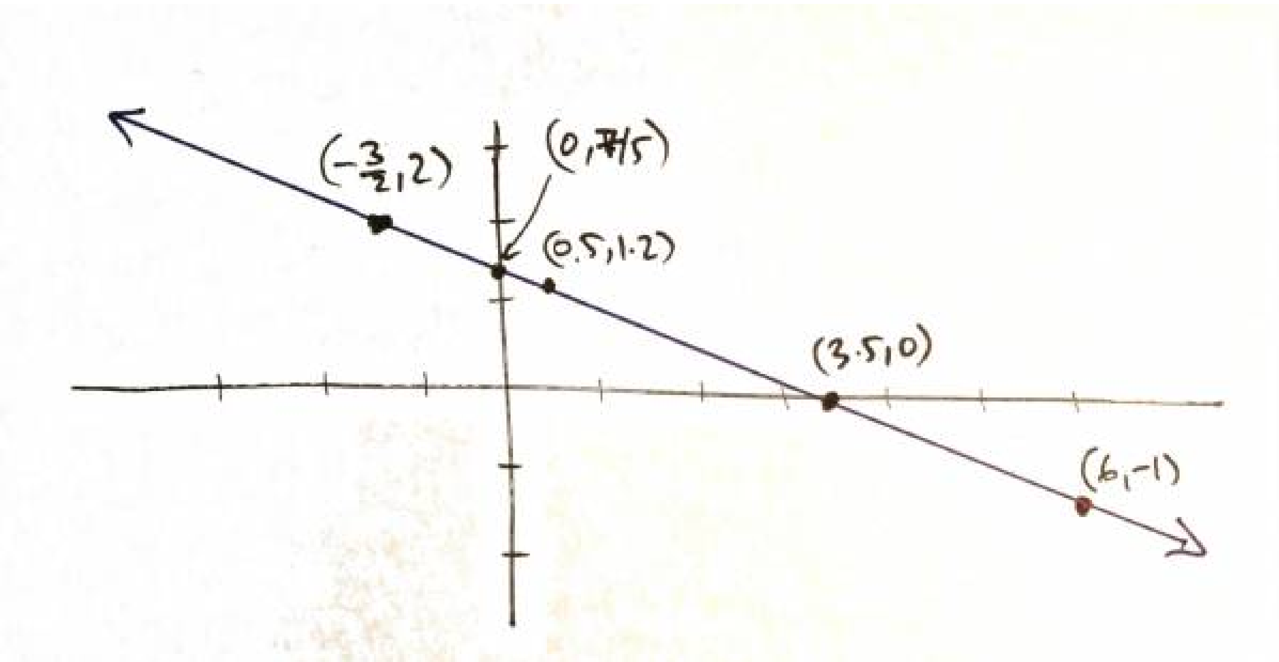
This is an equation in two *variables*, or *indeterminates*, x and y.

A solution of this equation is a pair of numbers $(a, b) \in \mathbb{R}^2$ so that if we replace x with a and replace y with b, then the equation becomes true.

In other words, so that $2a + 5b$ really is equal to 7.

- $(3, 1)$ is not a solution, because $2 \times 3 + 5 \times 1 \neq 7$
- $(1, 1)$ is a solution, because $2 \times 1 + 5 \times 1 = 7$
- Other solutions include $(0, \frac{7}{5})$, $(0.5, 1.2)$, $(6, -1)$, $(3.5, 0)$, $(-\frac{3}{2}, 2)$, …

We can't make a complete list of all solutions, since there are *infinitely many* solutions in $\mathbb{R}^2$. However, we can draw the set of all solutions as a subset of $\mathbb{R}^2$. This turns out to be a straight line:



We say that the equation $2x + 5y = 7$ is a linear equation in two variables.

Definition

If a, b, c are any fixed numbers, then equation

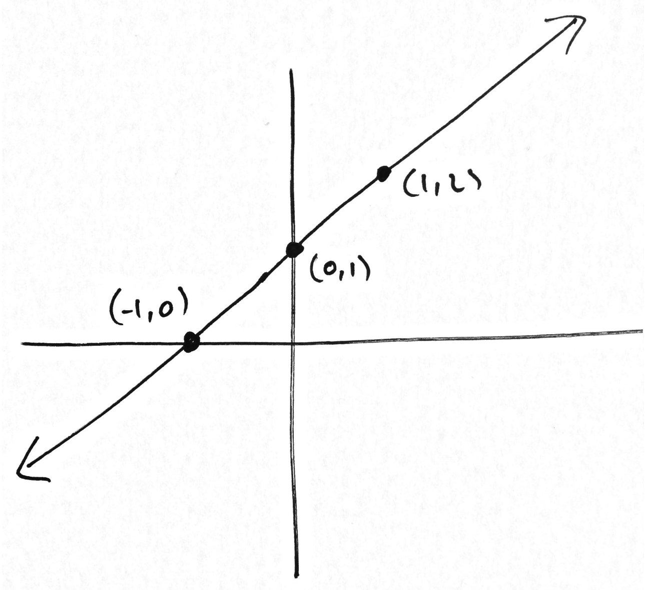
$$ax + by = c$$

is a **linear equation in two variables**.

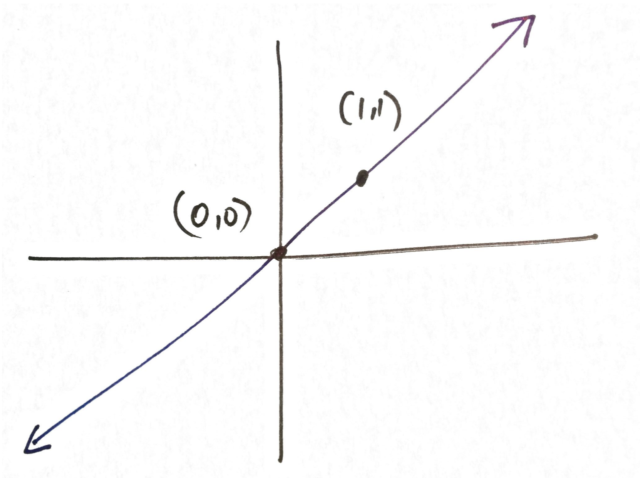
When you draw the set of all solutions of a linear equation in two variables, you always get a straight line in the x-y plane.

More examples of linear equations in two variables

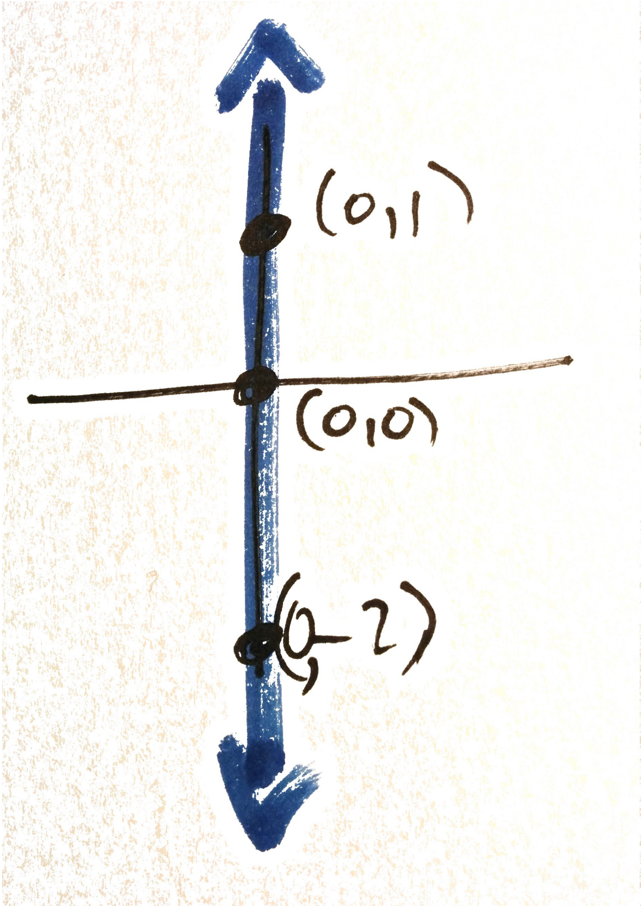
- $y - x = 1$



- $x - y = 0$



- $x = 0 \iff 1x + 0y = 0$


(0,1)
(0,0)
(0,-2)

Linear equations in 3 variables

Definition

If a, b, c, d are any fixed numbers, then equation

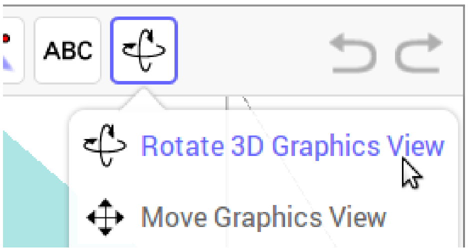
$$ax + by + cz = d$$

is a **linear equation in 3 variables**.

When you draw the set of all solutions of a linear equation in 3 variables, you always get a plane in 3-dimensional space, $\mathbb{R}^3$.

Examples

Note: you can view the examples below from different angles, by clicking the “Rotate 3D graphics view” button.



- $x + y + z = 1$

- $x + y = 1$ This may be viewed as a linear equation in 3 variables, since it is equivalent to $x + y + 0z = 1$.

- $z = 1$, viewed as the equation $0x + 0y + z = 1$

This plane is horizontal (parallel to the x-y plane).

Linear equations (in general)

A **linear equation** in m variables (where m is some natural number [https://en.wikipedia.org/wiki/natural number]) is an equation of the form

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

where $a_1, a_2, \ldots, a_m$ and b are fixed numbers (called **coefficients**) and $x_1, x_2, \ldots, x_m$ are variables.

Example

$$3x_1 + 5x_2 - 7x_3 + 11x_4 = 12$$

is a linear equation in 4 variables.

- A typical solution will be a point $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ so that $3x_1 + 5x_2 - 7x_3 + 11x_4$ really does equal 12.
- For example, $(-2, 0, -1, 1)$ is a solution.
- The set of all solutions is a 3-dimensional object in $\mathbb{R}^4$, called a hyperplane [https://en.wikipedia.org/wiki/hyperplane].
- Since we can't draw pictures in 4-dimensional space $\mathbb{R}^4$ we can't draw this set of solutions!

Systems of linear equations

A **system of linear equations** is just a list of several linear equations. By a solution of the system, we mean a common solution of each equation in the system.

Example

Find the line of intersection of the two planes

$$x + 3y + z = 5$$

and

$$2x + 7y + 4z = 17.$$

Just to get an idea of what's going on, here's a picture of the two planes:

To find the equation of the line of intersection, we must find the points which are solutions of *both* equations at the same time. Eliminating variables, we get

$$x = -16 + 5z, \quad y = 7 - 2z$$

which tells us that for any value of z, the point

$$(-16 + 5z, 7 - 2z, z)$$

is a typical point in the line of intersection.

Let's look at the example from the end of Lecture 2 more closely:

$$\begin{array}{rcrcrcrl} x & + & 3y & + & z & = & 5 & \qquad (1) \\ 2x & + & 7y & + & 4z & = & 17 & \qquad (2) \end{array}$$

We find the solutions of this system by applying operations to the system to make a new system, aiming to end up with a very simple sort of system where we can see the solutions easily.

First replace equation (2) with $(2) - 2 \times (1)$. We'll call the resulting equations (1) and (2) again, although of course we end up with a different system of linear equations:

$$\begin{array}{rcrcrcrl} x & + & 3y & + & z & = & 5 & \qquad (1) \\ & & y & + & 2z & = & 7 & \qquad (2) \end{array}$$

Now replace equation (1) with $(1) - 3 \times (2)$:

$$\begin{array}{rcrcrcrl} x & & & - & 5z & = & -16 & \qquad (1) \\ & & y & + & 2z & = & 7 & \qquad (2) \end{array}$$

Notice that we can now easily rearrange (1) to find x in terms of z, and we can rearrange (2) to find y in terms of z. Since z can take any value, we write $z = t$ where t is a "free parameter" (which means t can be any real number, or $t \in \mathbb{R}$).

$$\begin{aligned} x &= -16 + 5t \\ y &= 7 - 2t \\ z &= t, \qquad t \in \mathbb{R} \end{aligned}$$

We can also write this in so-called "vector form":

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -16 \\ 7 \\ 0 \end{bmatrix} + t \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \qquad t \in \mathbb{R}.$$

This is the equation of the line where the two planes described by the original equations (1) and (2) intersect.

Note for each different value of t, we get a different solutions (that is, a different point on the line of intersection). For example, setting $t = 0$ we see that $(-16, 7, 0)$ is a solution; setting $t = 1.5$, we see that $(-16 + 1.5 \times 5, 7 + 1.5 \times (-2), 1.5) = (-8.5, 4, 1.5)$ is another solution, and so on. This works for any value $t \in \mathbb{R}$, and every solution may be written in this way.

Observations

1. The operations we applied to the original linear system don't change the set of solutions. This is because each operation is reversible.
2. Writing out the variables x, y, z each time is unnecessary. If we erase the variables from the system

$$\begin{array}{rcrcrcrl} x & + & 3y & + & z & = & 5 & \qquad (1) \\ 2x & + & 7y & + & 4z & = & 17 & \qquad (2) \end{array}$$

and write all the numbers in a grid, or a **matrix**, we get:

$$\begin{bmatrix} 1 & 3 & 1 & 5 \\ 2 & 7 & 4 & 17 \end{bmatrix}$$

Notice that the first column corresponds to the x variable, the second to y, the third to z and the numbers in the final column are the right hand sides of the equations. Each row corresponds to one equation. So instead of performing operations on equations, we can perform operations on the rows of this matrix:

$$\begin{aligned} & \begin{bmatrix} 1 & 3 & 1 & 5 \\ 2 & 7 & 4 & 17 \end{bmatrix} \\ \xrightarrow{R2 \to R2 - 2 \times R1} & \begin{bmatrix} 1 & 3 & 1 & 5 \\ 0 & 1 & 2 & 7 \end{bmatrix} \\ \xrightarrow{R1 \to R1 - 3 \times R1} & \begin{bmatrix} 1 & 0 & -5 & -16 \\ 0 & 1 & 2 & 7 \end{bmatrix} \end{aligned}$$

Now we translate this back into equations to solve:

$$\begin{array}{ccccccc} x & & & - & 5z & = & -16 \qquad (1) \\ & & y & + & 2z & = & 7 \qquad (2) \end{array}$$

so

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -16 \\ 7 \\ 0 \end{bmatrix} + t \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \qquad t \in \mathbb{R}.$$

This sort of thing works in general: we can take any system of linear equations, write down a corresponding matrix, perform certain reversible operations on the rows of this matrix to get a new matrix, and then write down a new system of linear equations with the same solutions as the original system. If we do things in a sensible way then the new system will be easy to solve, so we'll be able to solve the original system (since the solution set is the same).

Let's give some terminology which will allow us to make this process clear.

The augmented matrix of a system of linear equations

Definition

Given a system of linear equations:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1m}x_m & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2m}x_m & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nm}x_m & = & b_n \end{array}$$

its **augmented matrix** is

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & b_n \end{bmatrix}.$$

The numbers in this matrix are called the **entries** of the matrix. We can be a bit more precise: the number in row i

and column j is called the (i, j) entry of the matrix.

Example

To find the augmented matrix of the linear system

$$\begin{aligned} 3x + 4y + 7z &= 2 \\ x + 3z &= 0 \\ y - 2z &= 5 \end{aligned}$$

notice that we can rewrite it as

$$\begin{aligned} 3x + 4y + 7z &= 2 \\ {\color{red}1}x + {\color{red}0}y + 3z &= 0 \\ {\color{red}0}x + {\color{red}1}y - 2z &= 5 \end{aligned}$$

so the augmented matrix is

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

- the $(2, 3)$ entry of this matrix is 3;
- the $(3, 2)$ entry is 1;
- the $(1, 4)$ entry is 2;
- the $(4, 1)$ entry is undefined (since this matrix does not have a 4th row).

Elementary operations on a system of linear equations

If we perform one of the following operations on a system of linear equations:

1. list the equations in a different order; or
2. multiply one of the equations by a non-zero real number; or
3. replace equation j by "equation $j + c \times$ (equation i)", where c is a non-zero real number,

then the new system will have exactly the same solutions as the original system. These are called **elementary operations** on the linear system.

Why do elementary operations leave the solutions of systems unchanged?

- we are doing the same thing to the left hand side and the right hand side of each equation, so any solution to the original system will also be a solution to the new system; and
- these operations are reversible, using operations of the same type, so any solution to the new system will also be a solution to the original system.

Elementary row operations on a matrix

Recall that when we form the augmented matrix of a linear system, each equation in the system becomes a row of the matrix. So we can translate the elementary operations on the linear system into corresponding operations on the rows of the matrix. We get three different types:

1. change the order of the rows of the matrix;
2. multiply one of the rows of the matrix by a non-zero real number;
3. replace row j by "row $j + c \times$ (row i)", where c is a non-zero real number and $i \neq j$.

The system of linear equations corresponding to these matrices will then have exactly the same solutions.

We call these operations **elementary row operations** or **EROs** on the matrix.

Example

Use EROs to find the intersection of the planes

$$\begin{aligned} 3x+4y+7z&=2\\ x+3z&=0\\ y-2z&=5 \end{aligned}$$

Solution 1

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

$$\xrightarrow{\text{reorder rows}} \begin{bmatrix} 1 & 0 & 3 & 0\\ 0 & 1 & -2 & 5\\ 3 & 4 & 7 & 2 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3 - 3R1} \begin{bmatrix} 1 & 0 & 3 & 0\\ 0 & 1 & -2 & 5\\ 0 & 4 & -2 & 2 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3 - 4R2} \begin{bmatrix} 1 & 0 & 3 & 0\\ 0 & 1 & -2 & 5\\ 0 & 0 & 6 & -18 \end{bmatrix}$$

$$\xrightarrow{R3 \to \frac{1}{6} R3} \begin{bmatrix} 1 & 0 & 3 & 0\\ 0 & 1 & -2 & 5\\ 0 & 0 & 1 & -3 \end{bmatrix}$$

So

- from the last row, we get $z=-3$
- from the second row, we get $y-2z=5$, so $y-2(-3)=5$, so $y=-1$
- from the first row, we get $x+3z=0$, so $x+3(-3)=0$, so $x=9$

The conclusion is that

$$\begin{bmatrix} x\\ y\\ z \end{bmatrix} = \begin{bmatrix} 9\\ -1\\ -3 \end{bmatrix}$$

is the only solution.

Example

Use EROs to find the intersection of the planes

$$\begin{aligned} 3x+4y+7z&=2\\ x+3z&=0\\ y-2z&=5 \end{aligned}$$

Solution 1

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

$$\xrightarrow{\text{reorder rows}} \begin{bmatrix} 1 & 0 & 3 & 0\\ 0 & 1 & -2 & 5\\ 3 & 4 & 7 & 2 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3 - 3R1} \begin{bmatrix} 1 & 0 & 3 & 0\\ 0 & 1 & -2 & 5\\ 0 & 4 & -2 & 2 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3 - 4R2} \begin{bmatrix} 1 & 0 & 3 & 0\\ 0 & 1 & -2 & 5\\ 0 & 0 & 6 & -18 \end{bmatrix}$$

$$\xrightarrow{R3 \to \frac{1}{6} R3} \begin{bmatrix} 1 & 0 & 3 & 0\\ 0 & 1 & -2 & 5\\ 0 & 0 & 1 & -3 \end{bmatrix}$$

So

- from the last row, we get $z=-3$
- from the second row, we get $y-2z=5$, so $y-2(-3)=5$, so $y=-1$
- from the first row, we get $x+3z=0$, so $x+3(-3)=0$, so $x=9$

The conclusion is that

$$\begin{bmatrix} x\\ y\\ z \end{bmatrix} = \begin{bmatrix} 9\\ -1\\ -3 \end{bmatrix}$$

is the only solution.

Solution 2

We start in the same way, but by performing more EROs we make the algebra at the end simpler.

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

$$\xrightarrow{\text{reorder rows}} \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 5 \\ 3 & 4 & 7 & 2 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3 - 3R1} \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 5 \\ 0 & 4 & -2 & 2 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3 - 4R2} \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 6 & -18 \end{bmatrix}$$

$$\xrightarrow{R3 \to \frac{1}{6}R3} \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 1 & -3 \end{bmatrix}$$

$$\xrightarrow{R2 \to R2 + 2R3} \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -3 \end{bmatrix}$$

$$\xrightarrow{R1 \to R1 - 3R3} \begin{bmatrix} 1 & 0 & 0 & 9 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -3 \end{bmatrix}$$

So

- from the last row, we get $z = -3$
- from the second row, we get $y = -1$
- from the first row, we get $x = 9$

The conclusion is again that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ -1 \\ -3 \end{bmatrix}$$

is the only solution.

Discussion

In both of these solutions we used EROs to transform the augmented matrix into a nice form.

- In solution 1, we ended up with the matrix $\begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 1 & -3 \end{bmatrix}$ which has a staircase pattern, with zeros below the staircase, and 1s just above the “steps” of the staircase. This is an example of a matrix in **row echelon form** (see below). We needed a bit of easy algebra, called **back substitution**, to finish off the solution. (Why is it called *echelon* form? It seems that this word has an archaic meaning [http://dictionary.reference.com/browse/echelon] which is relevant to the staircase-like pattern: “any structure or group of structures arranged in a steplike form.”)
- In solution 2, we ended up with the matrix $\begin{bmatrix} 1 & 0 & 0 & 9 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -3 \end{bmatrix}$ which has a staircase pattern with zeros below the

staircase and 1s just above the "steps" of the staircase, and the additional property that we only have zeros above the 1s on the steps. This is an example of a matrix in **reduced row echelon form** (see below). Finding the solution from this matrix needed no extra algebra.

Row echelon form and reduced row echelon form

Row echelon form (REF)

Definition

A row of a matrix is a **zero row** if it contains only zeros. For example, $[0\ 0\ 0\ 0\ 0]$ is a zero row.

A row of a matrix is **non-zero**, or a **non-zero row** if contains at least one entry that is not 0. For example $[0\ 0\ 3\ 0\ 0]$ is non-zero, and so is $[1\ 2\ 3\ 4\ \ -5]$.

Definition

The **leading entry** of a non-zero row of a matrix is the leftmost entry which is not 0.

For example, the leading entry of the row $[0\ 0\ 0\ 6\ 2\ 0\ 3\ 1\ 0]$ is 6.

Definition

A matrix is in **row echelon form**, or **REF**, if it has all of the following three properties:

1. The zero rows of the matrix (if any) are all at the bottom of the matrix.
2. In every non-zero row of the matrix, the leading entry is 1.
3. If row i and row $(i+1)$ are both non-zero, then the leading entry in row $(i+1)$ is to the right of the leading entry in row i.
   In other words, as you go down the rows, the leading entries must go to the right.

For example, $\begin{bmatrix} 1&2&3&4&5\\0&1&2&3&4\\0&0&1&2&3 \end{bmatrix}$ and $\begin{bmatrix} 1&2&3&4&5\\0&1&2&3&4\\0&0&1&2&3\\0&0&0&0&0 \end{bmatrix}$ are both in REF, but

1. $\begin{bmatrix} 1&2&3&4&5\\0&0&0&0&0\\0&0&1&2&3 \end{bmatrix}$ and $\begin{bmatrix} 1&2&3&4&5\\0&0&0&0&0\\0&0&1&2&3\\0&0&0&0&0 \end{bmatrix}$ are not in REF, since they each have a zero row which isn't at the bottom;
2. $\begin{bmatrix} 1&2&3&4&5\\0&2&3&4&1\\0&0&1&2&3 \end{bmatrix}$ is not in REF, since the leading entry on the second row isn't 1;
3. $\begin{bmatrix} 0&1&2&3&4\\1&2&3&4&5\\0&0&1&2&3 \end{bmatrix}$ is not in REF, since the leading entry in row 2 is not to the right of the leading entry in row 1.

Reduced row echelon form (RREF)

Definition

A matrix is in **reduced row echelon form** or **RREF** if it is in row echelon form (REF), so that

1. The zero rows of the matrix (if any) are all at the bottom of the matrix.
2. In every non-zero row of the matrix, the leading entry is 1.
3. If row i and row $(i+1)$ are both non-zero, then the leading entry in row $(i+1)$ is to the right of the leading entry in row i.
   In other words, as you go down the rows, the leading entries must go to the right.

and the matrix also has the property:

4. If a column contains the leading entry of a row, then every other entry in that column is 0.

For example,

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 3 & 4 & 5 \\ 0 & 1 & 0 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix}$$

are both in REF, but they are not in RREF because the red entries are non-zero and are in the same column as a leading entry (in blue).

On the other hand,

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

is in RREF.

Example

Use EROs to put the following matrix into RREF:

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

and solve the corresponding linear system.

Solution

$$\begin{aligned} & \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix} \\ \xrightarrow{R2 \to R2 - 2R3} & \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 0 & -1 & -2 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix} \\ \xrightarrow{R1 \to R1 - 3R3} & \begin{bmatrix} 1 & 2 & 0 & -2 & -4 \\ 0 & 1 & 0 & -1 & -2 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix} \\ \xrightarrow{R1 \to R1 - 2R2} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -2 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix} \end{aligned}$$

This matrix is in RREF. Write x_i for the variable corresponding to the ith column. The solution is

1. $x_4 = t$, a free parameter, i.e. $t \in \mathbb{R}$. This is because the 4th column does not contain a leading entry.
2. From row 3: $x_3 + 2t = 3$, so $x_3 = 3 - 2t$
3. From row 2: $x_2 - t = -2$, so $x_2 = -2 + t$
4. From row 1: $x_1 = 0$

So the solution is

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

(Geometrically, this is a line in 4-dimensional space $\mathbb{R}^4$).

Solving a system in REF or RREF

Given an augmented matrix in REF (or RREF), each column except the last column corresponds to a variable. These come in two types:

- **leading variables** are variables whose column contains the leading entry of some row;
- **free variables** are all the other variables.

Example

For the augmented matrix

$$\begin{bmatrix} 1 & 2 & 3 & 0 & 0 & 8 \\ 0 & 0 & 1 & 1 & 1 & 5 \\ 0 & 0 & 0 & 1 & 3 & 4 \end{bmatrix}$$

which is in REF, if we use the variables x_1, x_2, x_3, x_4, x_5 then

- x_1, x_3 and x_4 are leading variables, since the corresponding columns have a leading entry
- x_2 and x_5 are free variables, since the corresponding columns do not have a leading entry

To solve such a linear system, we use the following procedure:

1. assign a free parameter (a letter like $r, s, t, \dots$ representing some arbitrary real number) to each free variable
2. starting at the bottom of the matrix, write out each row and rearrange it to give an equation for its leading variable, substituting the other variables as needed.

In the example above, this gives:

1. x_2 and x_5 are free, so set $x_2 = s$ where $s \in \mathbb{R}$ and $x_5 = t$ where $t \in \mathbb{R}$
2. Working from the bottom:

$$x_4 + 3x_5 = 4 \implies x_4 = 4 - 3x_5 = 4 - 3t$$

$$x_3 + x_4 + x_5 = 5 \implies x_3 = 5 - x_4 - x_5 = 5 - (4 - 3t) - t = 1 + 2t$$

$$x_1 + 2x_2 + 3x_3 = 8 \implies x_1 = 8 - 2x_2 - 3x_3 = 8 - 2s - 3(1 + 2t) = 5 - 2s - 6t.$$

So

$$\begin{aligned} x_1 &= 5 - 2s - 6t \\ x_2 &= s \\ x_3 &= 1 + 2t \\ x_4 &= 4 - 3t \\ x_5 &= t \end{aligned}$$

where s and t are free parameters, i.e. $s, t \in \mathbb{R}$.

Writing this solution in vector form gives:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 1 \\ 4 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -6 \\ 0 \\ 2 \\ -3 \\ 1 \end{bmatrix}, \quad s, t \in \mathbb{R}.$$

Note that the solution set is a subset of $\mathbb{R}^5$, which is 5-dimensional space; and the solution set is 2-dimensional, because there are 2 free parameters.

Gaussian elimination

We've seen that putting a matrix into REF (or even better, in RREF) makes it easier to solve equations.

Aim: put any matrix into REF using EROs.

Algorithm

1. Re-order the rows so that the leftmost leading entry in the matrix is in the top row.
2. Divide all of the top row by its leading entry, so that this entry becomes a 1.
3. “Pivot about the leading 1”: subtract multiples of the top row from each row below so that all entries below the leading 1 in the top row become 0.
4. Go back to the start, ignoring the top row (until no rows remain, except possibly zero rows).

If we want, we can go further and put the matrix into RREF. First put it into REF as above, and then:

1. Look at the non-zero row nearest the bottom of the matrix
2. Pivot about the leading 1 in that row and use it to make zeros above: subtract multiples of that row from each row above
3. Move to the next row up, and go to step 2 (until no rows remain).

Example

Use Gaussian elimination to solve the linear system

$$\begin{aligned} 2x + y + 3z + 4w &= 27 \\ x + 2y + 3z + 2w &= 30 \\ x + y + 3z + w &= 25 \end{aligned}$$

Solution 1

We put the augmented matrix into REF:

$$\begin{bmatrix} 2 & 1 & 3 & 4 & 27 \\ 1 & 2 & 3 & 2 & 30 \\ 1 & 1 & 3 & 1 & 25 \end{bmatrix}$$

$$\xrightarrow{\text{reorder rows (to avoid division)}} \begin{bmatrix} 1 & 1 & 3 & 1 & 25 \\ 1 & 2 & 3 & 2 & 30 \\ 2 & 1 & 3 & 4 & 27 \end{bmatrix}$$

$$\xrightarrow{R2 \to R2 - R1 \text{ and } R3 \to R3 - 2R2} \begin{bmatrix} 1 & 1 & 3 & 1 & 25 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & -1 & -3 & 2 & -23 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3 + R2} \begin{bmatrix} 1 & 1 & 3 & 1 & 25 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & 0 & -3 & 3 & -18 \end{bmatrix}$$

$$\xrightarrow{R3 \to -\frac{1}{3} R3} \begin{bmatrix} 1 & 1 & 3 & 1 & 25 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & 0 & 1 & -1 & 6 \end{bmatrix}$$

This is in REF. There are leading entries in the columns for x, y, z but not for w, so $w = t$ is a free variable (where $t \in \mathbb{R}$). Now

$$z - w = 6 \implies z = 6 + w = 6 + t$$

$$y + w = 5 \implies y = 5 - w = 5 - t$$

$$x + y + 3z + w = 25 \implies x = 25 - y - 3z - w = 25 - (5 - t) - 3(6 + t) - t = 2 - 3t.$$

So

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \qquad t \in \mathbb{R}.$$

Solution 2

We put the augmented matrix into RREF.

$$\begin{bmatrix} 2 & 1 & 3 & 4 & 27 \\ 1 & 2 & 3 & 2 & 30 \\ 1 & 1 & 3 & 1 & 25 \end{bmatrix}$$

$$\xrightarrow{\text{do everything as above}} \begin{bmatrix} 1 & 1 & 3 & 1 & 25 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & 0 & 1 & -1 & 6 \end{bmatrix}$$

$$\xrightarrow{R1 \to R1 - 3R3} \begin{bmatrix} 1 & 1 & 0 & 4 & 7 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & 0 & 1 & -1 & 6 \end{bmatrix}$$

$$\xrightarrow{R1 \to R1 - R2} \begin{bmatrix} 1 & 0 & 0 & 3 & 2 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & 0 & 1 & -1 & 6 \end{bmatrix}$$

This is in RREF. There are leading entries in the columns for x, y, z but not for w, so $w = t$ is a free variable (where $t \in \mathbb{R}$). Now

$$z - w = 6 \implies z = 6 + w = 6 + t$$

$$y + w = 5 \implies y = 5 - w = 5 - t$$

$$x + 3w = 2 \implies x = 2 - 3w = 2 - 3t$$

So

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Example

A function $f(x)$ has the form

$$f(x) = ax^2 + bx + c$$

where a, b, c are constants. Given that $f(1) = 3$, $f(2) = 2$ and $f(3) = 4$, find $f(x)$.

Solution

$$\begin{aligned} f(1) = 3 &\implies a \cdot 1^2 + b \cdot 1 + c = 3 \implies a + b + c = 3 \\ f(2) = 2 &\implies a \cdot 2^2 + b \cdot 2 + c = 3 \implies 4a + 2b + c = 2 \\ f(3) = 4 &\implies a \cdot 3^2 + b \cdot 3 + c = 3 \implies 9a + 3b + c = 4 \end{aligned}$$

We get a system of three linear equations in the variables a, b, c:

$$\begin{aligned} a + b + c &= 3 \\ 4a + 2b + c &= 2 \\ 9a + 3b + c &= 4 \end{aligned}$$

Let's reduce the augmented matrix for this system to RREF.

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ 4 & 2 & 1 & 2 \\ 9 & 3 & 1 & 4 \end{bmatrix}$$

$$\xrightarrow{R2 \to R2-4R1 \text{ and } R3 \to R3-9R1} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & -3 & -10 \\ 0 & -6 & -8 & -23 \end{bmatrix}$$

$$\xrightarrow{R2 \to -\frac{1}{2} R2} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & \frac{3}{2} & 5 \\ 0 & -6 & -8 & -23 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3+6R2} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & \frac{3}{2} & 5 \\ 0 & 0 & 1 & 7 \end{bmatrix}$$

$$\xrightarrow{R1 \to R1-R3 \text{ and } R2 \to R2-\frac{3}{2} R3} \begin{bmatrix} 1 & 1 & 0 & -4 \\ 0 & 1 & 0 & -5.5 \\ 0 & 0 & 1 & 7 \end{bmatrix}$$

$$\xrightarrow{R1 \to R1-R2} \begin{bmatrix} 1 & 0 & 0 & 1.5 \\ 0 & 1 & 0 & -5.5 \\ 0 & 0 & 1 & 7 \end{bmatrix}$$

So $a = 1.5$, $b = -5.5$ and $c = 7$; so

$$f(x) = 1.5x^2 - 5.5x + 7.$$

Examples

Example 1

Solve the linear system

$$\begin{aligned} 2x + 4y - 2z &= 1 \\ x + y + z &= 4 \\ 2x + 5y + z &= 3 \end{aligned}$$

by transforming the augmented matrix into reduced row echelon form.

Solution

$$\begin{aligned}
& \begin{bmatrix} 3 & 4 & -2 & 1 \\ 1 & 1 & 1 & 4 \\ 2 & 5 & 1 & 3 \end{bmatrix} \\
\xrightarrow{R1 \leftrightarrow R2} & \begin{bmatrix} 1 & 1 & 1 & 4 \\ 3 & 4 & -2 & 1 \\ 2 & 5 & 1 & 3 \end{bmatrix} \\
\xrightarrow{R2 \to R2 - 3R1 \text{ and } R3 \to R3 - 2R1} & \begin{bmatrix} 1 & 1 & 1 & 4 \\ 0 & 1 & -5 & -11 \\ 0 & 3 & -1 & -5 \end{bmatrix} \\
\xrightarrow{R3 \to R3 - 3R1} & \begin{bmatrix} 1 & 1 & 1 & 4 \\ 0 & 1 & -5 & -11 \\ 0 & 0 & 14 & 28 \end{bmatrix} \\
\xrightarrow{R3 \to \frac{1}{14} R3} & \begin{bmatrix} 1 & 1 & 1 & 4 \\ 0 & 1 & -5 & -11 \\ 0 & 0 & 1 & 2 \end{bmatrix} \\
\xrightarrow{R1 \to R1 - R3 \text{ and } R2 \to R2 + 5R3} & \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \\
\xrightarrow{R1 \to R1 - R2} & \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}
\end{aligned}$$

The solution is $x = 3$, $y = -1$, $z = 2$. So there is a unique solution: just one point in $\mathbb{R}^3$, namely $(3, -1, 2)$.

Example 2

Solve the linear system

$$\begin{aligned} 3x + y - 2z + 4w &= 5 \\ x + z + w &= 2 \\ 4x + 2y - 6z + 6w &= 0 \end{aligned}$$

Solution

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

$$\xrightarrow{R1 \leftrightarrow R2} \begin{bmatrix} 1 & 0 & 1 & 1 & 2 \\ 3 & 1 & -2 & 4 & 5 \\ 4 & 2 & -6 & 6 & 0 \end{bmatrix}$$

$$\xrightarrow{R2 \to R2 - 3R1 \text{ and } R3 \to R3 - 4R1} \begin{bmatrix} 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & -5 & 1 & -1 \\ 0 & 2 & -10 & 2 & -8 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3 - 2R2} \begin{bmatrix} 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & -5 & 1 & -1 \\ 0 & 0 & 0 & 0 & -6 \end{bmatrix}$$

$$\xrightarrow{R3 \to -\frac{1}{6}R3} \begin{bmatrix} 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & -5 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

This is in REF. The last row corresponds to the equation

$$0 = 1$$

which clearly has no solution! We conclude that this system has no solutions, and hence the original linear system has no solutions either.

A linear system with no solutions is called **inconsistent**.

We can detect an inconsistent linear system, since whenever we apply EROs to put the augmented matrix into REF, we will get a row of the form $[0\ 0\ 0\ \ldots\ 0\ *]$ where $*$ is non-zero.

Example 3

For which value(s) of k does the following linear system have infinitely many solutions?

$$\begin{aligned} x + y + z &= 1 \\ x - z &= 5 \\ 2x + 3y + kz &= -2 \end{aligned}$$

Solution

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

$$\xrightarrow{R2 \to R2 - R1 \text{ and } R3 \to R3 - 2R1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & 4 \\ 0 & 1 & k-2 & -4 \end{bmatrix}$$

$$\xrightarrow{R2 \to -R2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & -4 \\ 0 & 1 & k-2 & -4 \end{bmatrix}$$

$$\xrightarrow{R3 \to R3 - R2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & -4 \\ 0 & 0 & k-4 & 0 \end{bmatrix}$$

If $k - 4 = 0$, then this matrix is in REF:

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

In this situation, z is a free variable (since there's no leading entry in the third column). For each value of z we get a different solution, so if $k - 4 = 0$, or equivalently, if $k = 4$, then there are infinitely many different solutions.

If $k - 4 \neq 0$, then we can divide the third row by $k - 4$ to get the REF:

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

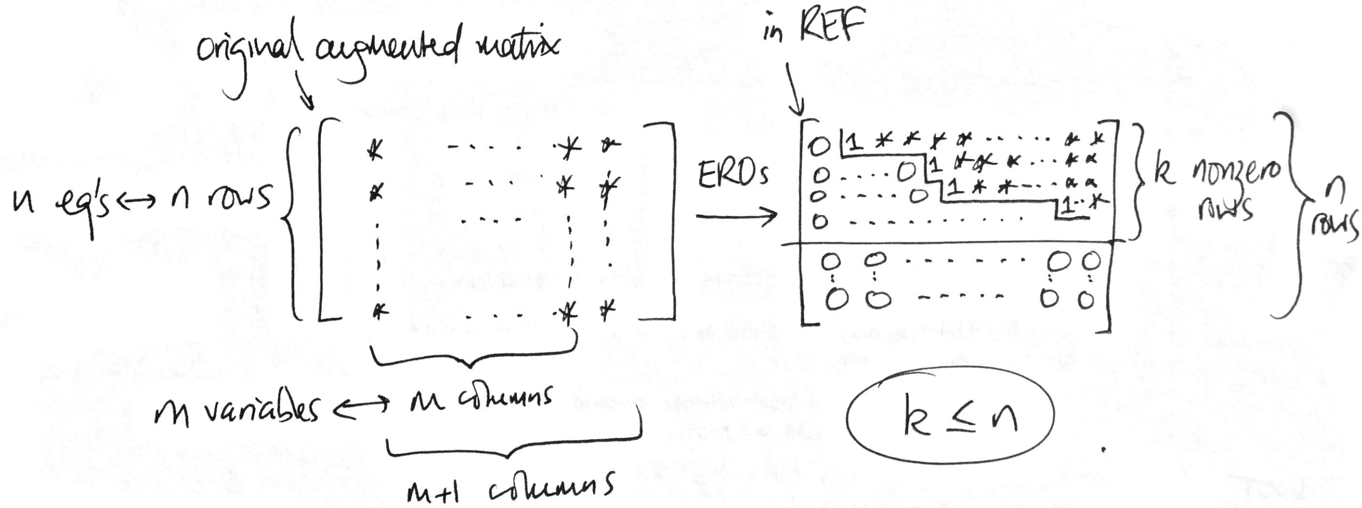
In this situtation, there are no free variables since x, y and z are all leading variables. So if $k-4 \neq 0$, or equivalently if $k \neq 4$, then there are no free variables so there is not an infinite number of solutions. (The only possibilities are that there is is a unique solution or that the system is inconsistent; and in this case you can check that there is a unique solution, although we don't need to know this to answer the question).

In conclusion, the system has infinitely many solutions if and only if $k = 4$.

Observations about Gaussian elimination

We know that we can apply EROs to any augmented matrix into REF.

Suppose the system has n equations and m variables, and let k be the number of non-zero rows in REF. Also suppose the system is consistent: then the REF has no row of the form $[0\ 0\ 0\ \ldots\ 1]$.



- $k \leq n$, because there are only n rows in the whole matrix
- k is precisely the number of leading variables. So k is no bigger m, the total number of variables; in symbols, we have $k \leq m$.
- All the other variables are free variables, so

$$m - k \text{ is the number of free variables.}$$

What does this tell us about the set of solutions? For example, how many solutions are there?

Observation 1: free variables and the number of solutions

For consistent systems, this shows that:

- *either* $k = m$;
  - so $m - k = 0$
  - there are no free variables
  - the system has one solution and no more
  - We say it has a **unique solution**.
- *or* $k < m$
  - so $m - k > 0$
  - there is at least one free variable
  - so the system has **infinitely many solutions** (one for each value of each free variable)
  - The number of free variables, $m - k$, is called the dimension [https://en.wikipedia.org/wiki/dimension] of the solution set.

Observation 2: systems with fewer equations than variables

For consistent systems where $n < m$ (fewer equations than variables):

- $k \le n < m$, so $k < m$.
- So there is at least one free variable.
- So in this situation we *always* have infinitely many solutions.

Chapter 2: The algebra of matrices

Definition

An $n \times m$ **matrix** is a grid of numbers with n rows and m columns:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}$$

The (i, j) **entry** of a matrix A is a_{ij}, the number in row i and column j of A.

Example

If $B = \begin{bmatrix} 99 & 3 & 5 \\ 7 & -20 & 14 \end{bmatrix}$, then B is a 2×3 matrix, and the $(1, 1)$ entry of B is $b_{11} = 99$, the $(1, 3)$ entry of B is $b_{13} = 5$, the $(2, 1)$ entry is $b_{21} = 7$, etc.

Examples

- $\begin{bmatrix} 3 \\ 2 \\ 4 \\ 0 \\ -1 \end{bmatrix}$ is a 5×1 matrix. A matrix like this with one column is called a **column vector**.
- $\begin{bmatrix} 3 & 2 & 4 & 0 & -1 \end{bmatrix}$ is a 1×5 matrix. A matrix like this with one row is called a **row vector**.

Even though the row matrix and the column matrix above have the same entries, they have a different "shape", or "size", so we must think of them has being different matrices. Let's give the definitions to make this precise.

Definition

Two matrices A and B have the **same size** if they have the same number of rows, and they have the same number of columns.

If two matrices do not have the same size, we say they have **different sizes**.

Definition

Two matrices A and B are said to be **equal** if both of the following conditions hold:

- A and B have the same size; and
- every entry of A is equal to the corresponding entry of B; in other words, for every (i, j) so that A and B have an (i, j) entry, we have $a_{ij} = b_{ij}$.

When A and B are equal matrices, we write $A = B$. Otherwise, we write $A \neq B$.

Examples

- $\begin{bmatrix} 3 \\ 2 \\ 4 \\ 0 \\ -1 \end{bmatrix} \neq \begin{bmatrix} 3 & 2 & 4 & 0 & -1 \end{bmatrix}$, since these matrices have different sizes: the first is 5×1 but the second is 1×5.
- $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$ since these matrices are not the same size.
- $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ because even though they have the same size, the $(2, 1)$ entries are different.
- If $\begin{bmatrix} 3x & 7y+2 \\ 8z-3 & w^2 \end{bmatrix} = \begin{bmatrix} 1 & 2z \\ \sqrt{2} & 9 \end{bmatrix}$ then we know that all the corresponding entries are equal, so we get four equations:

$$
\begin{aligned}
3x &= 1 \\
7y + 2 &= 2z \\
8z - 3 &= \sqrt{2} \\
w^2 &= 9
\end{aligned}
$$

Operations on matrices

We want to define operations on matrices: some (useful) ways of taking two matrices and making a new matrix.

Before we begin, a remark about 1×1 matrices. These are of the form $[a_{11}]$ where a_{11} is just a number. The square brackets $[\,]$ don't really matter here; they just keep the inside of a matrix in one place. So really: a 1×1 matrix is just a number. This means that special cases of the operations we define will be operations on ordinary numbers. You should check that in the special case when all the matrices involved are 1×1 matrices, the operations become the ordinary operations on numbers, so we are *generalising* the familiar operations (addition, subtraction, multiplication and so on) from numbers to matrices.

Matrix addition and subtraction

Definition of matrix addition

If A and B are matrices of the same size, then $A + B$ is defined to be the matrix with the same size as A and B so that the (i, j) entry of $A + B$ is $a_{ij} + b_{ij}$, for every i, j.

If A and B are matrices of different sizes, then $A + B$ is undefined.

Example

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

Example

$$\begin{bmatrix} 1 & 2 & -2 \\ 3 & 0 & 5 \end{bmatrix} + \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix} \text{ is undefined.}$$

Remarks

1. For any matrices A and B with the same size: $A + B = B + A$. We say that matrix addition is *commutative*.
2. For any matrices A, B and C with the same size: $(A + B) + C = A + (B + C)$. We say that matrix addition is *associative*.

Definition of the zero matrix

The $n \times m$ **zero matrix** is the $n \times m$ matrix so that every entry is 0. We write this as $0_{n\times m}$. So

$$0_{n\times m} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

where this matrix has n rows and m columns.

Exercise

Show that if A is any $n \times m$ matrix, then

$$0_{n\times m} + A = A = A + 0_{n\times m}.$$

Remember that when checking that matrices are equal, you have to check that they have the same size, and that all the entries are the same.

Definition of matrix subtraction

If A and B are matrices of the same size, then $A - B$ is defined to be the matrix with the same size as A and B so that the (i, j) entry of $A - B$ is $a_{ij} - b_{ij}$, for every i, j.

If A and B are matrices of different sizes, then $A - B$ is undefined.

Example

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

Example

$$\begin{bmatrix} 1 & 2 & -2 \\ 3 & 0 & 5 \end{bmatrix} - \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix} \text{ is undefined.}$$

Scalar multiplication

Definition of a scalar

In linear algebra, a **scalar** is just a fancy name for a number (in this course: a *real* number). The reason is that numbers are often used for scaling things up or down; for example, the scalar 3 is often used to scale things up by a factor of 3 (by multiplying by 3).

Definition of scalar multiplication of matrices

If c is a real number and A is an $n \times m$ matrix, then we define the matrix cA to be the $n \times m$ matrix given by multiplying every entry of A by c. In other words, the (i, j) entry of cA is $ca_{i,j}$.

Example

If $A = \begin{bmatrix} 1 & 0 & -3 \\ 3 & -4 & 1 \end{bmatrix}$, then $3A = \begin{bmatrix} 3 & 0 & -9 \\ 9 & -12 & 3 \end{bmatrix}$. In other words,

$$3\begin{bmatrix} 1 & 0 & -3 \\ 3 & -4 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & -9 \\ 9 & -12 & 3 \end{bmatrix}.$$

The negative of a matrix

We write $-A$ as a shorthand for $-1A$; so the (i, j) entry of $-A$ is $-a_{ij}$. For example,

$$-\begin{bmatrix} -1 & 0 & 3 \\ 3 & -4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -3 \\ -3 & 4 & -1 \end{bmatrix}.$$

Exercise

Prove that $A - B = A + (-B)$ for any matrices A and B of the same size.

Row-column multiplication

Definition of row-column multiplication

If $a = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$ is a $1 \times n$ row vector and $b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ is an $n \times 1$ column vector, then the **row-column product**, or simply the **product** of a and b is defined to be

$$ab = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

If we want to emphasize that we are multiplying in this way, we sometimes write $a \cdot b$ instead of ab.

The product ab of a $1 \times m$ row vector a with an $n \times 1$ column vector b is undefined if $m \neq n$.

Examples

- $\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot (-1) = 3 + (-2) = 1.$
- $\begin{bmatrix} 1 & 2 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ is not defined.
- $\begin{bmatrix} 2 & 3 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2x + 3y + 5z.$
- Generalising the previous example: if $a = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}$ and $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$, then $ax = a_1 x_1 + a_2 x_2 + \cdots + a_m x_m$. So we can write any linear equation $a_1 x_1 + a_2 x_2 + \cdots + a_m x_m = b$ as a shorter matrix equation: $ax = b$.

Matrix multiplication

This generalises row-column multiplication. The idea is that you build a new matrix from all possible row-column products. The formal definition will appear later, but here's an example:

$$\begin{bmatrix} 1 & 0 & 5 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \\ \begin{bmatrix} 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} & \begin{bmatrix} 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 26 & 32 \\ 14 & 18 \end{bmatrix}.$$

Definition of matrix multiplication

If A is an $n\times m$ matrix and B is an $m\times k$ matrix, then the product AB is the $n\times k$ matrix whose (i,j) entry is the row-column product of the ith row of A with the jth column of B. That is:

$$(AB)_{i,j}=\mathrm{row}_i(A)\cdot\mathrm{col}_j(B).$$

If we want to emphasize that we are multiplying matrices in this way, we might sometimes write $A\cdot B$ instead of AB.

If A is an $n\times m$ matrix and B is an $\ell\times k$ matrix with $m\neq\ell$, then the matrix product AB is undefined.

Examples

1. If $A=\begin{bmatrix}1&0&5\\2&-1&3\end{bmatrix}$ and $B=\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}$, then $AB=\begin{bmatrix}26&32\\14&18\end{bmatrix}$ and $BA=\begin{bmatrix}5&-2&11\\11&-4&27\\17&-6&43\end{bmatrix}$. Note that AB and BA are both defined, but $AB\neq BA$ since AB and BA don't even have the same size.
2. If $A=\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}$, $B=\begin{bmatrix}2&1&1\\1&2&0\\1&0&2\\2&2&1\end{bmatrix}$ and $C=\begin{bmatrix}1&3&0&7\\0&4&6&8\end{bmatrix}$, then A is 3×2, B is 4×3 and C is 2×4, so
   - AB, CA and BC don't exist (i.e., they are undefined);
   - AC exists and is 3×4;
   - BA exists and is 4×2; and
   - CB exists and is 2×2.
   - In particular, $AB\neq BA$ and $AC\neq CA$ and $BC\neq CB$, since in each case one of the matrices doesn't exist.
3. If $A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$ and $B=\begin{bmatrix}0&0\\1&0\end{bmatrix}$, then $AB=\begin{bmatrix}1&0\\0&0\end{bmatrix}$ and $BA=\begin{bmatrix}0&0\\0&1\end{bmatrix}$. So AB and BA are both defined and have the same size, but they are not equal matrices: $AB\neq BA$.
4. If $A=\begin{bmatrix}1&0\\0&0\end{bmatrix}$ and $B=\begin{bmatrix}0&0\\0&1\end{bmatrix}$, then $AB=\begin{bmatrix}0&0\\0&0\end{bmatrix}$ and $BA=\begin{bmatrix}0&0\\0&0\end{bmatrix}$. So $AB=BA$ in this case.
5. If $A=0_{n\times n}$ is the $n\times n$ zero matrix and B is any $n\times n$ matrix, then $AB=0_{n\times n}$ and $BA=0_{n\times n}$. So in this case, we do have $AB=BA$.
6. If $A=\begin{bmatrix}1&2\\3&4\end{bmatrix}$ and $B=\begin{bmatrix}7&10\\15&22\end{bmatrix}$, then $AB=\begin{bmatrix}37&54\\81&118\end{bmatrix}=BA$, so $AB=BA$ for these particular matrices A and B.
7. If $A=\begin{bmatrix}1&2\\3&4\end{bmatrix}$ and $B=\begin{bmatrix}6&10\\15&22\end{bmatrix}$, then $AB=\begin{bmatrix}36&54\\78&118\end{bmatrix}$ and $BA=\begin{bmatrix}36&52\\81&118\end{bmatrix}$, so $AB\neq BA$.

Commuting matrices

We say that matrices A and B **commute** if $AB=BA$.

Which matrices commute? Suppose A is an $n \times m$ matrix and B is an $\ell \times k$ matrix, and A and B commute, i.e., $AB = BA$.

- AB must be defined, so $m = \ell$
- BA must be defined, so $k = n$
- AB is an $n \times k$ matrix and BA is an $\ell \times n$ matrix. Since AB has the same size as BA, we must have $n = \ell$ and $k = m$.

Putting this together: we see that if A and B commute, then A and B must both be $n \times n$ matrices for some number n. In other words, they must be *square matrices of the same size*.

Examples 4 and 5 above show that for some square matrices A and B of the same size, it is true that A and B commute. On the other hand, examples 3 and 6 show that it's not true that square matrices of the same size must always commute.

Because it's not true in general that $AB = BA$, we say that **matrix multiplication is not commutative**.

Definition of the $n \times n$ identity matrix

The $n \times n$ **identity matrix** is the $n \times n$ matrix I_n with 1s in every diagonal entry (that is, in the (i, i) entry for every i between 1 and n), and 0s in every other entry. So

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Examples

1. $I_1 = [1]$
2. $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
3. $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
4. $I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, and so on!

Proposition: properties of I_n

1. $I_n A = A$ for any $n \times m$ matrix A;
2. $AI_m = A$ for any $n \times m$ matrix A; and
3. $I_n B = B = BI_n$ for any $n \times n$ matrix B. In particular, I_n commutes with every other square $n \times n$ matrix B.

Proof of the proposition

1. We want to show that $I_n A = A$ for any $n \times m$ matrix A. These matrices the same size, since I_n has size $n \times n$ and A has size $n \times m$, so $I_n A$ has size $n \times m$ by the definition of matrix multiplication, which is the same as the size of A.

Note that $\mathbf{row}_i(I_n) = [0\ 0\ \dots\ 0\ 1\ 0\ \dots\ 0]$, where the 1 is in the ith place, by definition of the identity matrix I_n; and the jth column of A is $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$. The (i,j) entry of I_nA is $\mathbf{row}_i(I_n) \cdot \mathbf{col}_j(A)$, by the definition of matrix multiplication, which is therefore

$$\begin{aligned}[0\ 0\ \dots\ 0\ 1\ 0\ \dots\ 0]\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix} &= 0a_{1j} + 0a_{2j} + \dots + 0a_{i-1,j} + 1a_{ij} + 0a_{i+1,j} + \dots + 0a_{nj} \\ &= a_{ij}.\end{aligned}$$

So the matrices I_nA and A have the same size, and the same (i,j) entries, for any (i,j). So $I_nA = A$.

Proof of the proposition, continued

2. To show that $AI_m = A$ for any $n \times m$ matrix A is similar to the first part of the proof; the details are left as an exercise.

3. If B is any $n \times n$ matrix, then $I_n B = B$ by part 1 and $BI_n = B$ by part 2, so $I_n B = B = BI_n$. In particular, $I_n B = BI_n$ so I_n commutes with B, for every square $n \times n$ matrix B. ■

Algebraic properties of matrix multiplication

The associative law

Proposition: associativity of matrix multiplication

Matrix multiplication is *associative*. This means that $(AB)C = A(BC)$ whenever A, B, C are matrices which can be multiplied together in this order.

We omit the proof, but this is not terribly difficult; it is a calculation in which you write down two formulae for the (i, j) entries of $(AB)C$ and $A(BC)$, and carefully check they are equal using the fact that if a, b, c are real numbers, then $(ab)c = a(bc)$.

Example

We saw above that $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ commutes with $B = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}$. We can explain why this is so using associativity. You can check that $B = AA$ (which we usually write as $B = A^2$). Hence, using associativity at $\stackrel{*}{=}$,

$$AB = A(AA) \stackrel{*}{=} (AA)A = BA.$$

The same argument for any square matrix A gives a proof of:

Proposition

If A is any square matrix, then A commutes with A^2.■

The powers of a square matrix A are defined by $A^1 = A$, and $A^{k+1} = A(A^k)$ for $k \in \mathbb{N}$. Using mathematical induction [https://en.wikipedia.org/wiki/mathematical induction], you can prove the following more general proposition.

Proposition: a square matrix commutes with its powers

If A is any square matrix and $k \in \mathbb{N}$, then A commutes with A^k.■

The distributive laws

Lemma: the distributive laws for row-column multiplication

1. If a is a $1 \times m$ row vector and b and c are $m \times 1$ column vectors, then $a \cdot (b + c) = a \cdot b + a \cdot c$.

2. If b and c are $1 \times m$ row vectors and a is an $m \times 1$ column vector, then $(b+c) \cdot a = b \cdot a + c \cdot a$.

The proof is an exercise (see tutorial worksheet 5).

Proposition: the distributive laws for matrix multiplication

If A is an $n \times m$ matrix and $k \in \mathbb{N}$, then:

1. $A(B+C) = AB + AC$ for any $m \times k$ matrices B and C; and
2. $(B+C)A = BA + CA$ for any $k \times n$ matrices B and C.

In other words, $A(B+C) = AB + AC$ whenever the matrix products make sense, and similarly $(B+C)A = BA + CA$ whenever this makes sense.

Proof

1. First note that

- B and C are both $m \times k$, so $B + C$ is $m \times k$ by the definition of matrix addition;
- A is $n \times m$ and $B + C$ is $m \times k$, so $A(B+C)$ is $m \times k$ by the definition of matrix multiplication;
- AB and AC are both $n \times k$ by the definition of matrix multiplication
- so $AB + AC$ is $n \times k$ by the definition of matrix addition.

So we have (rather long-windedly) checked that $A(B+C)$ and $AB + AC$ have the same size.

By the Lemma above, the row-column product has the property that

$$a \cdot (b+c) = a \cdot b + a \cdot c.$$

So the (i,j) entry of $A(B+C)$ is

$$\begin{aligned} \mathbf{row}_i(A) \cdot \mathbf{col}_j(B+C) &= \mathbf{row}_i(A) \cdot \big(\mathbf{col}_j(B) + \mathbf{col}_j(C)\big) \\ &= \mathbf{row}_i(A) \cdot \mathbf{col}_j(B) + \mathbf{row}_i(A) \cdot \mathbf{col}_j(C). \end{aligned}$$

On the other hand,

- the (i,j) entry of AB is $\mathbf{row}_i(A) \cdot \mathbf{col}_j(B)$; and
- the (i,j) entry of AC is $\mathbf{row}_i(A) \cdot \mathbf{col}_j(C)$;
- so the (i,j) entry of $AB + AC$ is also $\mathbf{row}_i(A) \cdot \mathbf{col}_j(B) + \mathbf{row}_i(A) \cdot \mathbf{col}_j(C)$.

So the entries of $A(B+C)$ and $AB + AC$ are all equal, so $A(B+C) = AB + AC$.

2. The proof is similar, and is left as an exercise.■

Matrix equations

We've seen that a single linear equation can be written using row-column multiplication. For example,

$$2x - 3y + z = 8$$

can be written as

$$\begin{bmatrix} 2 & -3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 8$$

or

$$a\vec{x} = 8$$

where $a = \begin{bmatrix} 2 & -3 & 1 \end{bmatrix}$ and $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

We can write a whole system of linear equations in a similar way, as a matrix equation using matrix multiplication. For example we can rewrite the linear system

$$\begin{aligned} 2x - 3y + z &= 8 \\ y - z &= 4 \\ x + y + z &= 0 \end{aligned}$$

as

$$\begin{bmatrix} 2 & -3 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 0 \end{bmatrix},$$

or

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

where $A = \begin{bmatrix} 2 & -3 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 8 \\ 4 \\ 0 \end{bmatrix}$. (We are writing the little arrow above the column vectors here because otherwise we might get confused between the $\vec{x}$: a column vector of variables, and x: just a single variable).

More generally, any linear system

$$\begin{matrix} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1m}x_m & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2m}x_m & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + \cdots + & a_{nm}x_m & = & b_n \end{matrix}$$

can be written in the form

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

where A is the $n \times m$ matrix, called the **coefficient matrix** of the linear system, whose (i, j) entry is a_{ij} (the number in front of x_j in the ith equation of the system) and $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

More generally still, we might want to solve a matrix equation like

$$AX = B$$

where A, X and B are matrices of any size, with A and B fixed matrices and X a matrix of unknown variables. Because of the definition of matrix multiplication, if A is $n \times m$, we need B to be $n \times k$ for some k, and then X must be $m \times k$, so we know the size of any solution X. But which $m \times k$ matrices X are solutions?

Example

If $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = 0_{2\times 3}$, then any solution X to $AX = B$ must be 2×3.

One solution is $X = 0_{2\times 3}$, since in this case we have $AX = A0_{2\times 3} = 0_{2\times 3}$.

However, this is not the only solution. For example, $X = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \end{bmatrix}$ is another solution, since in this case

$$AX = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0_{2\times 3}.$$

So from this example, we see that a matrix equation can have many solutions.

Invertibility

We've seen that solving matrix equations $AX = B$ is useful, since they generalise systems of linear equations.

How can we solve them?

Example

Take $A = \begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$, so we want to find all matrices X so that $AX = B$, or

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

Note that X must be a 2×2 matrix for this to work, by the definition of matrix multiplication. So one way to solve this is to write $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ and plug it in:

$$\begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} \iff \begin{bmatrix} 2x_{11} + 4x_{21} & 2x_{12} + 4x_{22} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$$

and then equate entries to get four linear equations:

$$\begin{aligned} 2x_{11} + 4x_{21} &= 3 \\ 2x_{12} + 4x_{22} &= 4 \\ x_{21} &= 5 \\ x_{22} &= 6 \end{aligned}$$

which we can solve in the usual way.

But this is a bit tedious! We will develop a slicker method by first thinking about solving ordinary equations $ax = b$ where a, x, b are all numbers, or if you like, 1×1 matrices.

Solving $ax = b$ and $AX = B$

If $a \neq 0$, then solving $ax = b$ where a, b, x are numbers is easy. We just divide both sides by a, or equivalently, we multiply both sides by $\frac{1}{a}$, to get the solution: $x = \frac{1}{a} \cdot b$.

Why does this work? If $x = \frac{1}{a} \cdot b$, then

$$\begin{aligned} ax &= a(\tfrac{1}{a} \cdot b) \\ &= (a \cdot \tfrac{1}{a})b \\ &= 1b \\ &= b \end{aligned}$$

so ax really is equal to b, and we do have a solution to $ax = b$.

What is special about $\frac{1}{a}$ which made this all work?

1. we have $a \cdot \frac{1}{a} = 1$,
2. and $1b = b$.

Now for an $n \times k$ matrix B, we know that the identity matrix I_n does the same sort of thing as 1 is doing in the relation $1b = b$: we have $I_n B = B$ for any $n \times k$ matrix B. So instead of $\frac{1}{a}$, we want to find a matrix C with the property: $AC = I_n$. In fact, because matrix multiplication is not commutative, we also require that $CA = I_n$. It's then easy to argue that $X = C \cdot B$ is a solution to $AX = B$, since

$$\begin{aligned} AX &= A(CB) \\ &= (AC)B \\ &= I_n B \\ &= B. \end{aligned}$$

Example revisited

If $A = \begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix}$, then the matrix $C = \begin{bmatrix} \frac{1}{2} & -2 \\ 0 & 1 \end{bmatrix}$ does have the property that

$$AC = I_2 = CA.$$

(You should check this!). So a solution to $AX = B$ where $B = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$ is

$$X = CB = \begin{bmatrix} \frac{1}{2} & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} -8.5 & -10 \\ 5 & 6 \end{bmatrix}.$$

Notice that having found the matrix C, then we can solve $AX = C$ easily for any 2×2 matrix C: the answer is

$X = CC$. This is quicker than having to solve four new linear equations using our more tedious method above.

Definition: invertible

An $n \times n$ matrix A is **invertible** if there exists an $n \times n$ matrix C so that

$$AC = I_n = CA.$$

The matrix C is called an **inverse** of A.

Examples

- $A = \begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix}$ is invertible, and the matrix $C = \begin{bmatrix} \frac{1}{2} & -2 \\ 0 & 1 \end{bmatrix}$ is an inverse of A
- a 1×1 matrix $A = [a]$ is invertible if and only if $a \neq 0$, and if $a \neq 0$ then an inverse of $A = [a]$ is $C = [\frac{1}{a}]$.
- I_n is invertible for any n, since $I_n \cdot I_n = I_n = I_n \cdot I_n$, so an inverse of I_n is I_n.
- $0_{n\times n}$ is not invertible for any n, since $0_{n\times n} \cdot C = 0_{n\times n}$ for any $n \times n$ matrix C, so $0_{n\times n} \cdot C \neq I_n$.
- $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is not invertible, since for any 2×2 matrix $C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ we have $AC = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ which is not equal to $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ since the $(2, 2)$ entries are not equal.
- $A = \begin{bmatrix} 1 & 2 \\ -3 & -6 \end{bmatrix}$ is not invertible. We'll see why later!

Proposition: uniqueness of the inverse

If A is an invertible $n \times n$ matrix, then A has a *unique* inverse.

Proof

Suppose C and C' are both inverses of A. Then $AC = I_n = CA$ and $AC' = I_n = C'A$. So

$$\begin{aligned} C &= CI_n \quad \text{by the properties of } I_n \\ &= C(AC') \quad \text{because } AC' = I_n \\ &= (CA)C' \quad \text{because matrix multiplication is associative} \\ &= I_nC' \quad \text{because } CA = I_n \\ &= C' \quad \text{by the properties of } I_n. \end{aligned}$$

So $C = C'$, whenever C and C' are inverses of A. So A has a unique inverse. ■

Definition/notation: A^{-1}

If A is an invertible $n \times n$ matrix, then the unique $n \times n$ matrix C with $AC = I_n = CA$ is called **the inverse of A**. If A is invertible, then we write A^{-1} to mean the (unique) inverse of A.

If a matrix A is not invertible, then A^{-1} does not exist.

Warning

If A is a matrix then $\frac{1}{A}$ doesn't make sense! You should never write this down. In particular, A^{-1} definitely doesn't mean $\frac{1}{A}$.

Similarly, you should **never** write down $\frac{A}{B}$ where A and B are matrices. This doesn't make sense either!

Examples revisited

- $A = \begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix}$ has $A^{-1} = \begin{bmatrix} \frac{1}{2} & -2 \\ 0 & 1 \end{bmatrix}$. In other words, $\begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & -2 \\ 0 & 1 \end{bmatrix}$.
- a 1×1 matrix $A = [a]$ with $a \neq 0$ has $[a]^{-1} = [\frac{1}{a}]$.
- $I_n^{-1} = I_n$.
- $0_{n\times n}^{-1}$ does not exist
- $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}^{-1}$ does not exist
- $\begin{bmatrix} 1 & 2 \\ -3 & -6 \end{bmatrix}^{-1}$ does not exist

Proposition: solving $AX = B$ when A is invertible

If A is an invertible $n \times n$ matrix and B is an $n \times k$ matrix, then the matrix equation

$$AX = B$$

has a unique solution: $X = A^{-1}B$.

Proof

First we check that $X = A^{-1}B$ really is a solution to $AX = B$. To see this, note that if $X = A^{-1}B$, then

$$\begin{aligned} AX &= A(A^{-1}B) \\ &= (AA^{-1})B \\ &= I_n B \\ &= B. \end{aligned}$$

Now we check that the solution is unique. If X and Y are both solutions, then $AX = B$ and $AY = B$, so

$$AX = AY.$$

Multiplying both sides on the left by A^{-1}, we get

$$A^{-1}AX = A^{-1}AY \implies I_n X = I_n Y \implies X = Y.$$

So any two solutions are equal, so $AX = B$ has a unique solution. ■

Corollary

If A is an $n \times n$ matrix and there is a non-zero $n \times m$ matrix K so that $AK = 0_{n\times m}$, then A is not invertible.

Proof

Since $A0_{n\times m} = 0_{n\times m}$ and $AK = 0_{n\times m}$, the equation $AX = 0_{n\times m}$ has (at least) two solutions: $X = 0_{n\times m}$ and $X = K$. Since K is non-zero, these two solutions are different.

So there is not a unique solution to $AX = B$, for B the zero matrix. If A was invertible, this would contradict the uniqueness statement of the last Proposition. So A cannot be invertible. ■

Examples

- We can now see why the matrix $A = \begin{bmatrix} 1 & 2 \\ -3 & -6 \end{bmatrix}$ is not invertible. If $X = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and $K = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, then K is non-zero, but $AK = 0_{2\times 1}$. So A is not invertible, by the Corollary.
- $A = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 5 & 7 \\ 3 & 6 & 9 \end{bmatrix}$ is not invertible, since $K = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ is non-zero and $AK = 0_{3\times 1}$.

2×2 matrices: determinants and invertibility

Question

Which 2×2 matrices are invertible? For the invertible matrices, can we find their inverse?

Lemma

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $J = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, then we have

$$AJ = \delta I_2 = JA$$

where $\delta = ad - bc$.

Proof

This is a calculation (done in the lectures; you should also check it yourself). ■

Definition: the determinant of a 2×2 matrix

The number $ad - bc$ is called the **determinant** of the 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. We write $\det(A) = ad - bc$ for this number.

Theorem: the determinant determines the invertibility (and inverse) of a 2×2 matrix

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix.

1. A is invertible if and only if $\det(A) \neq 0$.
2. If A is invertible, then $A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Proof

If $A = 0_{2\times 2}$, then $\det(A) = 0$ and A is not invertible. So the statement is true is this special case.

Now assume that $A \neq 0_{2\times 2}$ and let $J = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

By the previous lemma, we have

$$AJ = (\det(A))I_2 = JA.$$

If $\det(A) \neq 0$, then multiplying this equation through by the scalar $\frac{1}{\det(A)}$, we get

$$A\left(\frac{1}{\det(A)}J\right) = I_2 = \left(\frac{1}{\det(A)}J\right)A,$$

so if we write $B = \frac{1}{\det(A)}J$ to make this look simpler, then we obtain

$$AB = I_2 = BA,$$

so in this case A is invertible with inverse $B = \frac{1}{\det(A)}J = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

If $\det(A) = 0$, then $AJ = 0_{2\times 2}$ and $J \neq 0_{2\times 2}$ (since $A \neq 0_{2\times 2}$, and J is obtained from A by swapping two entries and multiplying the others by -1). Hence by the previous corollary, A is not invertible in this case. ■

Example

Let's solve the matrix equation $\begin{bmatrix} 1 & 5 \\ 3 & -2 \end{bmatrix} X = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$ for X.

Write $A = \begin{bmatrix} 1 & 5 \\ 3 & -2 \end{bmatrix}$. Then $\det(A) = 1(-2) - 5(3) = -2 - 15 = -17$ which isn't zero, so A is invertible. And $A^{-1} = \frac{1}{-17}\begin{bmatrix} -2 & -5 \\ -3 & 1 \end{bmatrix} = \frac{1}{17}\begin{bmatrix} 2 & 5 \\ 3 & -1 \end{bmatrix}$.

Hence the solution is $X = A^{-1}\begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \frac{1}{17}\begin{bmatrix} 2 & 5 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \frac{1}{17}\begin{bmatrix} 8 & 12 & 5 \\ 12 & 1 & -1 \end{bmatrix}$.

The transpose of a matrix

We defined this in tutorial sheet 4:

The transpose of an $n \times m$ matrix A is the $m \times n$ matrix A^T whose (i, j) entry is the (j, i) entry of A. In other words, to get A^T from A, you write the rows of A as columns, and vice versa; equivalently, you reflect A in its main diagonal.

For example, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

Exercise: simple properties of the transpose

Prove that for any matrix A:

- $(A^T)^T = A$; and
- $(A + B)^T = A^T + B^T$ if A and B are matrices of the same size; and
- $(cA)^T = c(A^T)$ for any scalar c.

In tutorial sheet 4, we proved:

Lemma: transposes and row-column multiplication

If a is a $1 \times m$ row vector and b is an $m \times 1$ column vector, then

$$ab = b^T a^T.$$

Observation: the transpose swaps rows with columns

Formally, for any matrix A and any i, j, we have

$$\begin{aligned} \mathrm{row}_i(A^T) &= \mathrm{col}_i(A)^T \\ \mathrm{col}_j(A^T) &= \mathrm{row}_j(A)^T. \end{aligned}$$

Theorem: the transpose reverses the order of matrix multiplication

If A and B are matrices and the matrix product AB is defined, then $B^T A^T$ is also defined. Moreover, in this case we have

$$(AB)^T = B^T A^T.$$

Proof

If AB is defined, then A is $n \times m$ and B is $m \times k$ for some n, m, k, so B^T is $k \times m$ and A^T is $m \times n$, so $B^T A^T$ is defined. Moreover, in this case $B^T A^T$ is an $k \times n$ matrix, and AB is an $n \times k$ matrix, so $(AB)^T$ is a $k \times n$ matrix. Hence $B^T A^T$ has the same size as $(AB)^T$. To show that they are equal, we calculate, using the fact that the transpose swaps rows with columns:

$$\begin{aligned}\text{the } (i,j) \text{ entry of } (AB)^T &= \text{the } (j,i) \text{ entry of } AB \\ &= \text{row}_j(A) \cdot \text{col}_i(B) \\ &= \text{col}_i(B)^T \cdot \text{row}_j(A)^T \quad \text{by the previous Lemma} \\ &= \text{row}_i(B^T) \cdot \text{col}_j(A^T) \quad \text{by the Observation} \\ &= \text{the } (i,j) \text{ entry of } B^T A^T\end{aligned}$$

Hence $(AB)^T = B^T A^T$. ■

Determinants of $n \times n$ matrices

Given any $n \times n$ matrix A, it is possible to define a number $\det(A)$ (as a formula using the entries of A) so that

$$A \text{ is invertible} \iff \det(A) \neq 0.$$

1. If A is a 1×1 matrix, say $A = [a]$, then we just define $\det[a] = a$.
2. If A is a 2×2 matrix, say $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then we've seen that $\det(A) = ad - bc$.
3. If A is a 3×3 matrix, say $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, then it turns out that $\det(A) = aei - afh + bfg - bdi + cdh - ceg$.
4. If A is a 4×4 matrix, then the formula for $\det(A)$ is more complicated still, with 24 terms.
5. If A is a 5×5 matrix, then the formula for $\det(A)$ has 120 terms.

Trying to memorise a formula in every case (or even in the 3×3 case!) isn't convenient unless we understand it somehow. We will approach this is several steps.

Step 1: minors

Definition

If A is an $n \times n$ matrix, then the (i,j) **minor** of A is defined to be the determinant of the $(n-1) \times (n-1)$ matrix formed by removing row i and column j from A. We will write this number as M_{ij}.

Examples

- If $A = \begin{bmatrix} 3 & 5 \\ -4 & 7 \end{bmatrix}$, then $M_{11} = \det[7] = 7$, $M_{12} = \det[-4] = -4$, $M_{21} = 5$, and $M_{22} = 3$.
- If $A = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 11 & 12 & 13 \end{bmatrix}$, then $M_{23} = \det \begin{bmatrix} 1 & 2 \\ 11 & 12 \end{bmatrix} = 1 \cdot 12 - 2 \cdot 11 = -10$ and $M_{32} = \det \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix} = -12$.

Step 2: cofactors

Definition

The (i,j) **cofactor** of an $n \times n$ matrix A is $(-1)^{i+j} M_{ij}$, where M_{ij} is the (i,j) minor of A.

Note that $(-1)^{i+j}$ is $+1$ or -1, and can looked up in the matrix of signs: $\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$. This matrix starts with a $+$ in the $(1,1)$ entry (corresponding to $(-1)^{1+1} = (-1)^2 = +1$) and the signs then alternate.

Examples

- If $A = \begin{bmatrix} 3 & 5 \\ -4 & 7 \end{bmatrix}$, then $C_{11} = +M_{11} = \det[7] = 7$, $C_{12} = -M_{12} = -\det[-4] = 4$, $C_{21} = -5$, and $C_{22} = 3$.
- If $A = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 11 & 12 & 13 \end{bmatrix}$, then $C_{23} = -M_{23} = -(-10) = 10$ and $C_{33} = +M_{33} = \det \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix} = -6$.

Step 3: the determinant of a 3×3 matrix using Laplace expansion along the first row

Definition

If $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ is a 3×3 matrix, then

$$\det A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}.$$

Here C_{ij} are the cofactors of A.

This formula is called the Laplace expansion of $\det A$ along the first row, since a_{11}, a_{12} and a_{13} make up the first row of A.

Example

$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 11 & 12 & 13 \end{bmatrix} &= 1 \cdot C_{11} + 2C_{12} + 3C_{13} \\ &= 1 \cdot (+M_{11}) + 2 \cdot (-M_{12}) + 3 \cdot (+M_{13}) \\ &= M_{11} - 2M_{12} + 3M_{13} \\ &= \det \begin{bmatrix} 8 & 9 \\ 12 & 13 \end{bmatrix} - 2\det \begin{bmatrix} 7 & 9 \\ 11 & 13 \end{bmatrix} + 3\det \begin{bmatrix} 7 & 8 \\ 11 & 12 \end{bmatrix} \\ &= (8 \cdot 13 - 9 \cdot 12) - 2(7 \cdot 13 - 9 \cdot 11) + 3(7 \cdot 12 - 8 \cdot 11) \\ &= -4 - 2(-8) + 3(-4) \\ &= -4 + 16 - 12 \\ &= 0. \end{aligned}$$

From this, we can conclude that $\begin{bmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 11 & 12 & 13 \end{bmatrix}$ is *not* invertible.

Notation

To save having to write $\det$ all the time, we sometimes write the entries of a matrix inside vertical bars $|\ |$ to mean the determinant of that matrix. Using this notation (and doing a few steps in our heads), we can rewrite the previous example as:

$$\begin{aligned}\begin{vmatrix}1&2&3\\7&8&9\\11&12&13\end{vmatrix}&=1\begin{vmatrix}8&9\\12&13\end{vmatrix}-2\begin{vmatrix}7&9\\11&13\end{vmatrix}+3\begin{vmatrix}7&8\\11&12\end{vmatrix}\\&=-4-2(-8)+3(-4)\\&=0.\end{aligned}$$

Step 4: the determinant of an $n\times n$ matrix

Definition

If $A=\begin{bmatrix}a_{11}&a_{12}&\dots&a_{1n}\\\vdots&&&\vdots\\a_{n1}&a_{n2}&\dots&a_{nn}\end{bmatrix}$ is an $n\times n$ matrix, then

$$\det A=a_{11}C_{11}+a_{12}C_{12}+\dots+a_{1n}C_{1n}.$$

Here C_{ij} are the cofactors of A.

This formula is called the Laplace expansion of $\det A$ along the first row, since $a_{11},a_{12},\dots,a_{1n}$ make up the first row of A.

Example

$$\begin{aligned}\begin{vmatrix}1&0&2&3\\0&2&1&-1\\2&0&0&1\\3&0&4&2\end{vmatrix}&=1\begin{vmatrix}2&1&-1\\0&0&1\\0&4&2\end{vmatrix}-0\begin{vmatrix}0&1&-1\\2&0&1\\3&4&2\end{vmatrix}+2\begin{vmatrix}0&2&-1\\2&0&1\\3&0&2\end{vmatrix}-3\begin{vmatrix}0&2&1\\2&0&0\\3&0&4\end{vmatrix}\\&=1\left(2\begin{vmatrix}0&1\\4&2\end{vmatrix}-1\begin{vmatrix}0&1\\0&2\end{vmatrix}-1\begin{vmatrix}0&0\\0&4\end{vmatrix}\right)-0+2\left(0-2\begin{vmatrix}2&1\\3&2\end{vmatrix}-1\begin{vmatrix}2&0\\3&0\end{vmatrix}\right)-3\left(0-2\begin{vmatrix}2&0\\3&4\end{vmatrix}+1\begin{vmatrix}2&0\\3&0\end{vmatrix}\right)\\&=1(2(-4)-0-0)+2(-2(1)-0)-3(-2(8)+0)\\&=-8-4+48\\&=36.\end{aligned}$$

Theorem: Laplace expansion along any row or column gives the determinant

1. For any fixed i: $\det(A)=a_{i1}C_{i1}+a_{i2}C_{i2}+\dots+a_{in}C_{in}$ (Laplace expansion along row i)
2. For any fixed j: $\det(A)=a_{1j}C_{1j}+a_{2j}C_{2j}+\dots+a_{nj}C_{nj}$ (Laplace expansion along column j)

Example

We can make life easier by choosing expansion rows or columns with lots of zeros, if possible. Let's redo the previous example with this in mind:

$$\begin{aligned}\begin{vmatrix}1&0&2&3\\0&2&1&-1\\2&0&0&1\\3&0&4&2\end{vmatrix}&=-0+2\begin{vmatrix}1&2&3\\2&0&1\\3&4&2\end{vmatrix}-0+0\\&=2\left(-2\begin{vmatrix}2&3\\4&2\end{vmatrix}+0-1\begin{vmatrix}1&2\\3&4\end{vmatrix}\right)\\&=2(-2(-8)-(-2))\\&=36.\end{aligned}$$

Definition: upper triangular matrices

An $n\times n$ matrix A is **upper triangular** if all the entries below the main diagonal are zero.

Definition: diagonal matrices

An $n \times n$ matrix A is **diagonal** if the only non-zero entries are on its main diagonal.

Corollary: the determinant of upper triangular matrices and diagonal matrices

1. The determinant of an upper triangular $n \times n$ matrix is the product of its diagonal entries: $\det(A) = a_{11}a_{22}\dots a_{nn}$.
2. The determinant of an $n \times n$ diagonal matrix is the product of its diagonal entries: $\det(A) = a_{11}a_{22}\dots a_{nn}$.

Proof

1. This is true for $n = 1$, trivially. For $n > 1$, assume inductively that it is true for $(n-1) \times (n-1)$ matrices and use the Laplace expansion of an upper triangular $n \times n$ matrix A along the first column of A to see that $\det(A) = a_{11}C_{11} + 0 + \dots + 0 = a_{11}C_{11}$. Now C_{11} is the determinant of the $(n-1) \times (n-1)$ matrix formed by removing the first row and and column of A, and this matrix is upper triangular with diagonal entries $a_{22}, a_{33}, \dots, a_{nn}$. By our inductive assumption, we have $C_{11} = a_{22}a_{33}\dots a_{nn}$. So $\det(A) = a_{11}C_{11} = a_{11}a_{22}a_{33}\dots a_{nn}$ as desired.
2. Any diagonal matrix is upper triangular, so this is a special case of statement 1. ■

Examples

1. For any n, we have $\det(I_n) = 1 \cdot 1 \cdots 1 = 1$.
2. For any n, we have $\det(5I_n) = 5^n$.
3. $\begin{vmatrix} 1 & 9 & 43 & 23434 & 4 & 132 \\ 0 & 3 & 43 & 2 & -1423 & -12 \\ 0 & 0 & 7 & 19 & 23 & 132 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -903 \\ 0 & 0 & 0 & 0 & 0 & 6 \end{vmatrix} = 1 \cdot 3 \cdot 7 \cdot 2 \cdot (-1) \cdot 6 = 252.$

Theorem: important properties of the determinant

Let A be an $n \times n$ matrix.

1. A is invertible if and only if $\det(A) \neq 0$.
2. $\det(A^T) = \det(A)$
3. If B is another $n \times n$ matrix, then $\det(AB) = \det(A)\det(B)$

Theorem: row/column operations and determinants

Let A be an $n \times n$ matrix, let c be a scalar and let $i \neq j$.

$A_{Ri \to x}$ means A but with row i replaced by x.

1. If $i \neq j$, then $\det(A_{Ri \leftrightarrow Rj}) = -\det(A)$ (swapping two rows changes the sign of det).
2. $\det(A_{Ri \to cRi}) = c\det(A)$ (scaling one row scales $\det(A)$ in the same way)
3. $\det(A_{Ri \to Ri + cRj}) = \det(A)$ (adding a multiple of one row to another row doesn't change $\det(A)$)

- Also, these properties all hold if you change “row” into “column” throughout.

Corollary

If an $n \times n$ matrix A has two equal rows (or columns), then $\det(A) = 0$, and A is not invertible.

Proof

If A has two equal rows, row i and row j, then $A = A_{Ri \leftrightarrow Rj}$ So $\det(A) = \det(A_{Ri \leftrightarrow Rj}) = -\det(A)$, so $2\det(A) = 0$, so $\det(A) = 0$.

If A has two equal columns, then A^T has two equal rows, so $\det(A) = \det(A^T) = 0$.

In either case, $\det(A) = 0$. So A is not invertible.■

Examples

- Swapping two rows changes the sign, so $\begin{vmatrix} 0 & 0 & 2 \\ 0 & 3 & 0 \\ 4 & 0 & 0 \end{vmatrix} = -\begin{vmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{vmatrix} = -4 \cdot 3 \cdot 2 = -24$.
- Multiplying a row or a column by a constant multiplies the determinant by that constant, so

$$\begin{aligned}
\begin{vmatrix} 2 & 4 & 6 & 10 \\ 5 & 0 & 0 & -10 \\ 9 & 0 & 81 & 99 \\ 1 & 2 & 3 & 4 \end{vmatrix} &= 2\begin{vmatrix} 1 & 2 & 3 & 5 \\ 5 & 0 & 0 & -10 \\ 9 & 0 & 81 & 99 \\ 1 & 2 & 3 & 4 \end{vmatrix} \\
&= 2 \cdot 5 \begin{vmatrix} 1 & 2 & 3 & 5 \\ 1 & 0 & 0 & -2 \\ 9 & 0 & 81 & 99 \\ 1 & 2 & 3 & 4 \end{vmatrix} \\
&= 2 \cdot 5 \cdot 9 \begin{vmatrix} 1 & 2 & 3 & 5 \\ 1 & 0 & 0 & -2 \\ 1 & 0 & 9 & 11 \\ 1 & 2 & 3 & 4 \end{vmatrix} \\
&= 2 \cdot 5 \cdot 9 \cdot 2 \begin{vmatrix} 1 & 1 & 3 & 5 \\ 1 & 0 & 0 & -2 \\ 1 & 0 & 9 & 11 \\ 1 & 1 & 3 & 4 \end{vmatrix} \\
&= 2 \cdot 5 \cdot 9 \cdot 2 \cdot 3 \begin{vmatrix} 1 & 1 & 1 & 5 \\ 1 & 0 & 0 & -2 \\ 1 & 0 & 3 & 11 \\ 1 & 1 & 1 & 4 \end{vmatrix}.
\end{aligned}$$

- $\det(A_{R1 \to R1 - R4}) = \det(A)$, so

$$\begin{vmatrix} 1 & 1 & 1 & 5 \\ 1 & 0 & 0 & -2 \\ 1 & 0 & 3 & 11 \\ 1 & 1 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & -2 \\ 1 & 0 & 3 & 11 \\ 1 & 1 & 1 & 4 \end{vmatrix} = -1 \begin{vmatrix} 1 & 0 & 0 \\ 1 & 0 & 3 \\ 1 & 1 & 1 \end{vmatrix} + 0$$
$$= - \begin{vmatrix} 0 & 3 \\ 1 & 1 \end{vmatrix} = -(-3) = 3.$$

- Hence

$$\begin{vmatrix} 2 & 4 & 6 & 10 \\ 5 & 0 & 0 & -10 \\ 9 & 0 & 81 & 99 \\ 1 & 2 & 3 & 4 \end{vmatrix} = 2 \cdot 5 \cdot 9 \cdot 2 \cdot 3 \begin{vmatrix} 1 & 1 & 1 & 5 \\ 1 & 0 & 0 & -2 \\ 1 & 0 & 3 & 11 \\ 1 & 1 & 1 & 4 \end{vmatrix}$$
$$= 2 \cdot 5 \cdot 9 \cdot 2 \cdot 3 \cdot 3 = 1620.$$

Corollary

If $\mathrm{row}_j(A) = c \cdot \mathrm{row}_i(A)$ for some $i \neq j$ and some $c \in \mathbb{R}$, then $\det(A) = 0$.

Proof

Note that $\mathrm{row}_i(A) - c \cdot \mathrm{row}_j(A) = 0$. So $A_{Ri \to Ri - c\,Rj}$ has a zero row, and by Laplace expansion along this row we obtain $\det(A_{Ri \to Ri - c\,Rj}) = 0$. So $\det(A) = \det(A_{Ri \to Ri - c\,Rj}) = 0. \blacksquare$

The effect of EROs on the determinant

We have now seen the effect of each of the three types of ERO on the determinant of a matrix:

1. swapping two rows of the matrix multiplies the determinant by -1. By swapping rows repeatedly, we are able to shuffle the rows in an arbitrary fashion, and the determinant will either remain unchanged (if we used an even number of swaps) or be multiplied by -1 (if we used an odd number of swaps).
2. multiplying one of the rows of the matrix by $c \in \mathbb{R}$ multiplies the determinant by c; and
3. replacing row j by "row $j + c \times$ (row i)", where c is a non-zero real number and $i \neq j$ does not change the determinant.

Moreover, since $\det(A) = \det(A^T)$, this all applies equally to columns instead of rows.

We can use EROs to put a matrix into upper triangular form, and then finding the determinant is easy: just multiply the diagonal entries together. We just have to keep track of how the determinant is changed by the EROs of types 1 and 2.

Example: using EROs to find the determinant

$$
\begin{vmatrix} 1 & 3 & 1 & 3 \\ 4 & 8 & 0 & 12 \\ 0 & 1 & 3 & 6 \\ 2 & 2 & 1 & 6 \end{vmatrix} = 4 \begin{vmatrix} 1 & 3 & 1 & 3 \\ 1 & 2 & 0 & 3 \\ 0 & 1 & 3 & 6 \\ 2 & 2 & 1 & 6 \end{vmatrix}
$$

$$
\begin{aligned}
&= 4 \cdot 3 \begin{vmatrix} 1 & 3 & 1 & 1 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 3 & 2 \\ 2 & 2 & 1 & 2 \end{vmatrix} \\
&= 12 \begin{vmatrix} 1 & 3 & 1 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & -4 & -1 & -0 \end{vmatrix} \\
&= -12 \begin{vmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & -1 & -1 & 0 \\ 0 & -4 & -1 & 0 \end{vmatrix} \\
&= -12 \begin{vmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 11 & 8 \end{vmatrix} \\
&= -12 \begin{vmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & -3 \end{vmatrix} \\
&= -12(1)(1)(2)(-3) = 72.
\end{aligned}
$$

Finding the inverse of an invertible $n \times n$ matrix

Definition: the adjoint of a square matrix

Let A be an $n \times n$ matrix. Recall that C_{ij} is the (i, j) cofactor of A. The **matrix of cofactors** of A is the $n \times n$ matrix C whose (i, j) entry is C_{ij}.

The **adjoint** of A is the $n \times n$ matrix $J = C^T$, the transpose of the matrix of cofactors.

Example: $n = 2$

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $C = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$, so the adjoint of A is $J = C^T = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Recall that $AJ = (\det A)I_2 = JA$; we calculated this earlier when we looked at the inverse of a 2×2 matrix. Hence for a 2×2 matrix A, if $\det A \neq 0$, then $A^{-1} = \frac{1}{\det A} J$.

Example: $n = 3$

If $A = \begin{bmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{bmatrix}$, then the matrix of signs is $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$, so

$$C = \begin{bmatrix} \begin{vmatrix} -4 & 3 \\ 4 & -2 \end{vmatrix} & -\begin{vmatrix} -2 & 3 \\ 5 & -2 \end{vmatrix} & \begin{vmatrix} -2 & -4 \\ 5 & 4 \end{vmatrix} \\ -\begin{vmatrix} 1 & 0 \\ 4 & -2 \end{vmatrix} & \begin{vmatrix} 3 & 0 \\ 5 & -2 \end{vmatrix} & -\begin{vmatrix} 3 & 1 \\ 5 & 4 \end{vmatrix} \\ \begin{vmatrix} 1 & 0 \\ -4 & 3 \end{vmatrix} & -\begin{vmatrix} 3 & 0 \\ -2 & 3 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ -2 & -4 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -4 & 11 & 12 \\ 2 & -6 & -7 \\ 3 & -9 & -10 \end{bmatrix}$$

so the adjoint of A is

$$J = C^T = \begin{bmatrix} -4 & 2 & 3 \\ 11 & -6 & -9 \\ 12 & -7 & -10 \end{bmatrix}.$$

Example: $n = 3$

If $A = \begin{bmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{bmatrix}$, then the matrix of signs is $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$, so

$$C = \begin{bmatrix} \begin{vmatrix} -4 & 3 \\ 4 & -2 \end{vmatrix} & -\begin{vmatrix} -2 & 3 \\ 5 & -2 \end{vmatrix} & \begin{vmatrix} -2 & -4 \\ 5 & 4 \end{vmatrix} \\ -\begin{vmatrix} 1 & 0 \\ 4 & -2 \end{vmatrix} & \begin{vmatrix} 3 & 0 \\ 5 & -2 \end{vmatrix} & -\begin{vmatrix} 3 & 1 \\ 5 & 4 \end{vmatrix} \\ \begin{vmatrix} 1 & 0 \\ -4 & 3 \end{vmatrix} & -\begin{vmatrix} 3 & 0 \\ -2 & 3 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ -2 & -4 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -4 & 11 & 12 \\ 2 & -6 & -7 \\ 3 & -9 & -10 \end{bmatrix}$$

so the adjoint of A is

$$J = C^T = \begin{bmatrix} -4 & 2 & 3 \\ 11 & -6 & -9 \\ 12 & -7 & -10 \end{bmatrix}.$$

Observe that $AJ = \begin{bmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{bmatrix} \begin{bmatrix} -4 & 2 & 3 \\ 11 & -6 & -9 \\ 12 & -7 & -10 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -1 \cdot I_3$, and
$JA = \begin{bmatrix} -4 & 2 & 3 \\ 11 & -6 & -9 \\ 12 & -7 & -10 \end{bmatrix} \begin{bmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -1 \cdot I_3$; and $\det(A) = -1$.

This is an illustration of the following theorem, whose proof is omitted:

Theorem: key property of the adjoint of a square matrix

If A is any $n \times n$ matrix and J is its adjoint, then $AJ = (\det A)I_n = JA$.

Corollary: a formula for the inverse of a square matrix

If A is any $n \times n$ matrix with $\det(A) \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{\det A} J$$

where J is the adjoint of A.

Proof

Divide the equation $AJ = (\det A)I_n = JA$ by $\det A$. ■

Example

If again we take $A = \begin{bmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{bmatrix}$, then $J = \begin{bmatrix} -4 & 2 & 3 \\ 11 & -6 & -9 \\ 12 & -7 & -10 \end{bmatrix}$ and $\det(A) = -1$, so A is invertible and $A^{-1} = \frac{1}{-1} J = -J = \begin{bmatrix} 4 & -2 & -3 \\ -11 & 6 & 9 \\ -12 & 7 & 10 \end{bmatrix}$.

A more efficient way to find A^{-1}

Given an $n \times n$ matrix A, form the $n \times 2n$ matrix

$$\left[\begin{array}{c|c} A & I_n \end{array}\right]$$

and use EROs to put this matrix into RREF. One of two things can happen:

- Either you get a row of the form $[0\ 0\ \dots\ 0 \mid *\ \ *\ \ \dots\ *]$ which starts with n zeros. You can then conclude that A is not invertible.
- Or you end up with a matrix of the form $\left[\begin{array}{c|c} I_n & B \end{array}\right]$ for some $n \times n$ matrix B. You can then conclude that A is invertible, and $A^{-1} = B$.

Examples

- Consider $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$.

$$\left[\begin{array}{c|c} A & I_2 \end{array}\right] = \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 6 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R2 \to R2 - 2R1} \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right]$$

  Conclusion: A is not invertible.

- Consider $A = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$.

$$\left[\begin{array}{c|c} A & I_2 \end{array}\right] = \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R2 \to R2 - 2R1} \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{array}\right]$$

$$\xrightarrow{R1 \to R1 - 3R1} \left[\begin{array}{cc|cc} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{array}\right]$$

  Conclusion: A is invertible and $A^{-1} = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}$.

- Consider $A = \begin{bmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{bmatrix}$.

$$\left[\begin{array}{c|c} A & I_3 \end{array}\right] = \left[\begin{array}{ccc|ccc} 3 & 1 & 0 & 1 & 0 & 0 \\ -2 & -4 & 3 & 0 & 1 & 0 \\ 5 & 4 & -2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R1 \to R1 + R2} \left[\begin{array}{ccc|ccc} 1 & -3 & 3 & 1 & 1 & 0 \\ -2 & -4 & 3 & 0 & 1 & 0 \\ 5 & 4 & -2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R2 \to R2 + 2R1,\ R3 \to R3 - 5R1} \left[\begin{array}{ccc|ccc} 1 & -3 & 3 & 1 & 1 & 0 \\ 0 & -10 & 9 & 2 & 3 & 0 \\ 0 & 19 & -17 & -5 & -5 & 1 \end{array}\right]$$

$$\xrightarrow{R3 \leftrightarrow R2} \left[\begin{array}{ccc|ccc} 1 & -3 & 3 & 1 & 1 & 0 \\ 0 & 19 & -17 & -5 & -5 & 1 \\ 0 & -10 & 9 & 2 & 3 & 0 \end{array}\right]$$

$$\xrightarrow{R2 \to R2 + 2R3} \left[\begin{array}{ccc|ccc} 1 & -3 & 3 & 1 & 1 & 0 \\ 0 & -1 & 1 & -1 & 1 & 1 \\ 0 & -10 & 9 & 2 & 3 & 0 \end{array}\right]$$

$$\xrightarrow{R1 \to R1 + 3R2,\ R3 \to R3 - 10R2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -2 & 3 \\ 0 & -1 & 1 & -1 & 1 & 1 \\ 0 & 0 & -1 & 12 & -7 & -10 \end{array}\right]$$

$$\xrightarrow{\overline{R2 \to R2 + R3}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -2 & 3 \\ 0 & -1 & 0 & 11 & -6 & -9 \\ 0 & 0 & -1 & 12 & -7 & -10 \end{array}\right]$$

$$\xrightarrow{R2 \to -R2,\ R3 \to -R3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -2 & 3 \\ 0 & 1 & 0 & -11 & 6 & 9 \\ 0 & 0 & 1 & -12 & 7 & 10 \end{array}\right]$$

Conclusion: A is invertible, and $A^{-1} = \begin{bmatrix} 4 & -2 & 3 \\ -11 & 6 & 9 \\ -12 & 7 & 10 \end{bmatrix}$.

Chapter 3: Vectors and geometry

Recall that a 2×1 column vector such as $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ is a pair of numbers written in a column. We are also used to writing points in the plane $\mathbb{R}^2$ as a pair of numbersl; for example $(4, 3)$ is the point obtained by starting from the origin, and moving 4 units to the right and 3 units up.

We think of a (column) vector like $\vec{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ as an instruction to move 4 units to the right and 3 units up. This movement is called “translation by $\vec{v}$”.

Examples

The vector $\vec{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ moves:

- $(0,0)$ to $(4,3)$
- $(-2,6)$ to $(2,9)$
- (x,y) to $(x+4,y+3)$.

It is convenient to not be too fussy about the difference between a point like $(4,3)$ and the vector $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$. If we agree to write points as column vectors, then we can perform algebra (addition, subtraction, scalar multiplication) as discussed in Chapter 2, using points and column vectors.

For example, we could rewrite the examples above by saying that $\vec{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ moves:

- $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ to $\begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$
- $\begin{bmatrix} -2 \\ 6 \end{bmatrix}$ to $\begin{bmatrix} -2 \\ 6 \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 9 \end{bmatrix}$
- $\begin{bmatrix} x \\ y \end{bmatrix}$ to $\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} x+4 \\ y+3 \end{bmatrix}$.

More generally: a column vector $\vec{v}$ moves a point $\vec{x}$ to $\vec{x} + \vec{v}$.

Example

Which vector moves the point $A = (-1, 3)$ to $B = (5, -4)$?

Answer: we need a vector $\vec{v}$ with $A + \vec{v} = B$, so $\vec{v} = B - A = \begin{bmatrix} 5 \\ -4 \end{bmatrix} - \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ -7 \end{bmatrix}$. We write $\vec{AB} = \begin{bmatrix} 6 \\ -7 \end{bmatrix}$, since this is the vector which moves A to B.

Definition of $\vec{AB}$

If A and B are any points in $\mathbb{R}^n$, then the vector $\vec{AB}$ is defined by

$$\vec{AB} = B - A$$

(where on the right hand side, we interpret the points as column vectors so we can subtract them to get a column vector).

Thus $\vec{AB}$ is the vector which moves the point A to the point B.

Example

In $\mathbb{R}^3$, the points $A = (3, -4, 5)$ and $B = (11, 6, -2)$ have $\vec{AB} = \begin{bmatrix} 11 \\ 6 \\ -2 \end{bmatrix} - \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix} = \begin{bmatrix} 8 \\ 10 \\ -7 \end{bmatrix}$.

The uses of vectors

Vectors are used in geometry and science to represent quantities with both a **magnitude** (size/length) and a **direction**. For example:

- displacements (in geometry)
- velocities
- forces

Recall that a column vector moves points. Its magnitude, or length, is how far it moves points.

Definition: the length of a vector

If $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ is a column vector in $\mathbb{R}^n$, then its **magnitude**, or **length**, or **norm**, is the number

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

Examples

- $\left\|\begin{bmatrix}4\\3\end{bmatrix}\right\| = \sqrt{4^2+3^2} = \sqrt{16+9} = \sqrt{25} = 5$
- $\left\|\begin{bmatrix}1\\0\\-2\\3\end{bmatrix}\right\| = \sqrt{1^2+0^2+(-2)^2+3^2} = \sqrt{1+0+4+9} = \sqrt{14}$

Exercise

Prove that if $c \in \mathbb{R}$ is a scalar and $\vec{v}$ is a vector in $\mathbb{R}^n$, then

$$\|c\vec{v}\| = |c|\ \|\vec{v}\|.$$

That is, multiplying a vector by a scalar c scales its length by $|c|$, the absolute value of c.

Remark

$\|\vec{AB}\|$ is the distance from point A to point B, since this is the length of vector which takes point A to point B.

Examples

- The distance from $A=(1,2)$ to $B=(-3,4)$ is $\|\vec{AB}\| = \left\|\begin{bmatrix}-3\\4\end{bmatrix} - \begin{bmatrix}1\\2\end{bmatrix}\right\| = \left\|\begin{bmatrix}-4\\2\end{bmatrix}\right\| = \sqrt{(-4)^2+2^2} = \sqrt{20} = 2\sqrt{5}$.
- The length of the main diagonal of the unit cube in $\mathbb{R}^3$ is the distance between $0=(0,0,0)$ and $A=(1,1,1)$, which is $\|\vec{0A}\| = \left\|\begin{bmatrix}1\\1\\1\end{bmatrix}\right\| = \sqrt{1^2+1^2+1^2} = \sqrt{3}$.

Scalar multiplication and direction

Multiplying a vector by a scalar changes its length, but doesn't change its direction.

Definition: unit vectors

A **unit vector** is a vector $\vec{v}$ with $\|\vec{v}\| = 1$.

Proposition: finding a unit vector in the same direction as a given vector

If $\vec{v}$ is a non-zero vector, then $\vec{w} = \frac{1}{\|\vec{v}\|}\,\vec{v}$ is a unit vector (in the same direction as $\vec{v}$).

Proof

Using the formula $\|c\vec{v}\| = |c|\ \|\vec{v}\|$ and the fact that $\|\vec{v}\| > 0$, we have

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

So $\vec{w}$ is a unit vector, and since it's scalar multiple of $\vec{v}$, it's in the same direction as $\vec{v}$. ■

Example

What is unit vector in the same direction as $\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$?

We have $||\vec{v}|| = \sqrt{1^2 + 2^2} = \sqrt{5}$, so the proposition tells us that is $\vec{w} = \frac{1}{||\vec{v}||}\vec{v} = \frac{1}{\sqrt{5}}\vec{v} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$ is a unit vector in the same direction as $\vec{v}$.

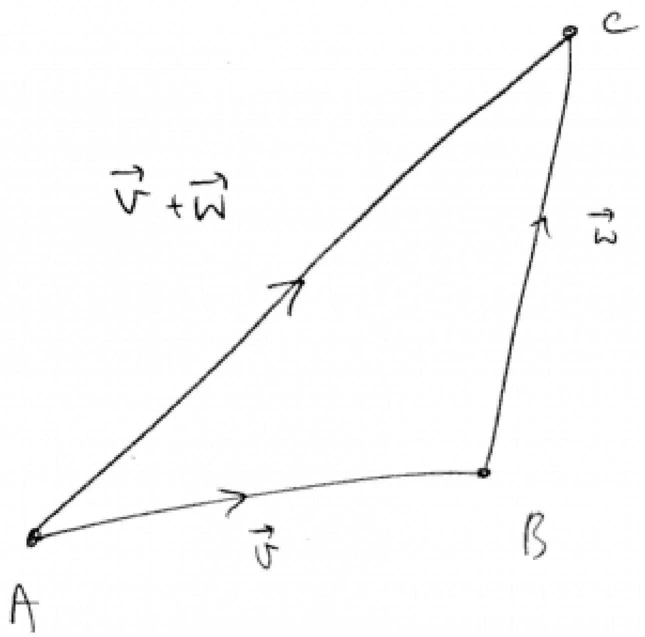
Addition of vectors

If $\vec{v} = \vec{AB}$, then $\vec{v}$ moves A to B, so $A + \vec{v} = B$.

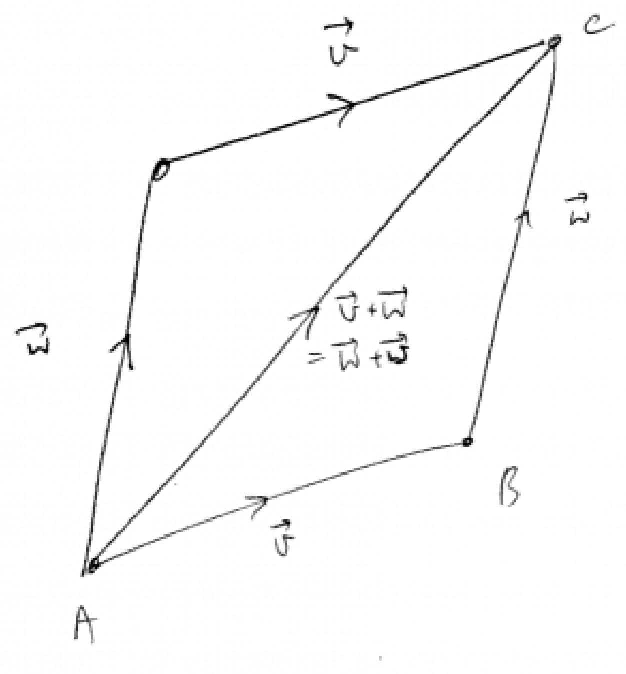
If $\vec{w} = \vec{BC}$, then $\vec{w}$ moves B to C, so $B + \vec{w} = C$.

What about $\vec{v} + \vec{w}$? We have $A + \vec{v} + \vec{w} = B + \vec{w} = C$. So $\vec{v} + \vec{w} = \vec{AC}$.

This gives us the **triangle law for vector addition**: $\vec{v}$, $\vec{w}$ and $\vec{v} + \vec{w}$ may be arranged to form a triangle:



We get another triangle by starting at A and translating first by $\vec{w}$ and then by $\vec{v}$; the other side of this triangle is $\vec{w} + \vec{v}$. But we know that $\vec{v} + \vec{w} = \vec{w} + \vec{v}$! So we can put these two triangles together to get the **parallelogram law for vector addition**:



C
$\vec{v}$
$\vec{w}$
$\vec{w}$
$\vec{v}+\vec{w}$
$=\vec{w}+\vec{v}$
B
$\vec{v}$
A

The dot product

Definition of the dot product

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 two (column) vectors in $\mathbb{R}^n$. The **dot product** of $\vec v$ and $\vec w$ is the real number $\vec v \cdot \vec w$ given by

$$\vec v \cdot \vec w = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n.$$

In other words, $\vec v \cdot \vec w$ is the row-column product $(\vec v)^T \vec w$ of the transpose of $\vec v$ with $\vec w$.

Note that while $\vec v$ and $\vec w$ are vectors, their dot product $\vec v \cdot \vec w$ is a scalar.

Example

If $\vec v = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ and $\vec w = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$, then $\vec v \cdot \vec w = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ -7 \end{bmatrix} = 3(4) + 5(-7) = -23$.

Properties of the dot product

For any vectors $\vec v$, $\vec w$ and $\vec u$ in $\mathbb{R}^n$, and any scalar $c \in \mathbb{R}$:

1. $\vec v \cdot \vec w = \vec w \cdot \vec v$ (the dot product is commutative)
2. $\vec u \cdot (\vec v + \vec w) = \vec u \cdot \vec v + \vec u \cdot \vec w$
3. $(c\vec v) \cdot \vec w = c(\vec v \cdot \vec w)$
4. $\vec v \cdot \vec v = \|\vec v\|^2 \geq 0$, and $\vec v \cdot \vec v = 0 \iff \vec v = 0_{n\times 1}$

The proofs of these properties are exercises.

Angles and the dot product

Theorem: the relationship between angle and the dot product

If $\vec v$ and $\vec w$ are non-zero vectors in $\mathbb{R}^n$, then

$$\vec v \cdot \vec w = \|\vec v\| \, \|\vec w\| \, \cos\theta$$

where θ is the angle between $\vec v$ and $\vec w$.

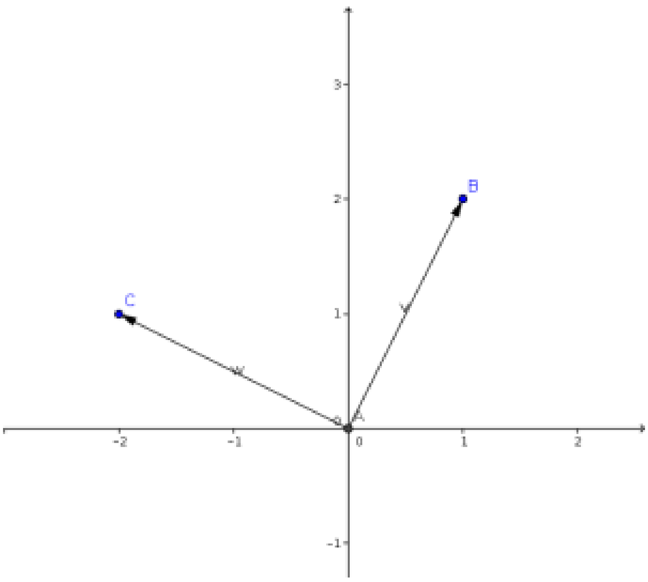
The proof will be given soon, but for now here is an example.

Example

If $\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, then $\vec{v} \cdot \vec{w} = 1(-2) + 2(1) = -2 + 2 = 0$. On the other hand, we have $\|\vec{v}\| = \sqrt{5} = \|\vec{w}\|$, so the angle θ between $\vec{v}$ and $\vec{w}$ satisfies

$$0 = \vec{v} \cdot \vec{w} = \sqrt{5} \times \sqrt{5} \times \cos\theta$$

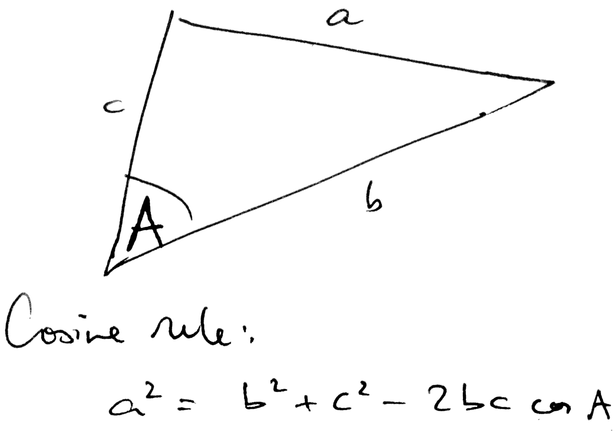
so $5\cos\theta = 0$, so $\cos\theta = 0$, so $\theta = \pi/2$ or $\theta = 3\pi/2$ (measuring angles in radians). This tells us that the angle between $\vec{v}$ and $\vec{w}$ is a right angle. We say that these vectors are **orthogonal**. We can draw a convincing picture which indicates that these vectors are indeed at right angles:



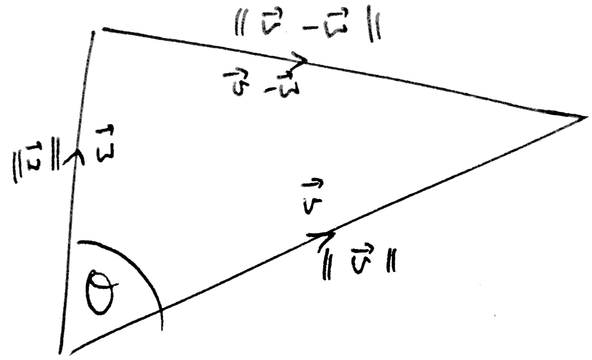
Proof of the Theorem

We wish to show that $\vec{v} \cdot \vec{w} = \|\vec{v}\| \, \|\vec{w}\| \cos\theta$ where θ is the angle between $\vec{v}$ and $\vec{w}$.

Recall the cosine rule [https://en.wikipedia.org/wiki/cosine rule]:



Consider a triangle with two sides $\vec{v}$ and $\vec{w}$. By the triangle rule for vector addition, the third side $\vec{x}$ has $\vec{w} + \vec{x} = \vec{v}$, so $\vec{x} = \vec{v} - \vec{w}$:



Applying the cosine rule gives

$$\|\vec{v}-\vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\|\vec{v}\|\,\|\vec{w}\|\,\cos\theta.$$

On the other hand, we know that $\|\vec{x}\|^2 = \vec{x}\cdot\vec{x}$, so

$$\begin{aligned}\|\vec{v}-\vec{w}\|^2 &= (\vec{v}-\vec{w})\cdot(\vec{v}-\vec{w})\\ &= \vec{v}\cdot\vec{v}+\vec{w}\cdot\vec{w}-\vec{w}\cdot\vec{v}-\vec{v}\cdot\vec{w}\\ &= \|\vec{v}\|^2+\|\vec{w}\|^2-2\vec{v}\cdot\vec{w}.\end{aligned}$$

So

$$\|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\|\vec{v}\|\,\|\vec{w}\|\,\cos\theta = \|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\vec{v}\cdot\vec{w}\cos\theta.$$

Subtracting $\|\vec{v}\|^2 + \|\vec{w}\|^2$ from both sides and dividing by -2 gives $\vec{v}\cdot\vec{w} = \|\vec{v}\|\,\|\vec{w}\|\cos\theta$. ■

Corollary

If $\vec{v}$ and $\vec{w}$ are non-zero vectors and θ is the angle between them, then $\cos\theta = \dfrac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\,\|\vec{w}\|}$.

Corollary

If $\vec{v}$ and $\vec{w}$ are non-zero vectors with $\vec{v}\cdot\vec{w}=0$, then $\vec{v}$ and $\vec{w}$ are **orthogonal**: they are at right-angles.

Examples

1. The angle θ between $\begin{bmatrix}1\\2\end{bmatrix}$ and $\begin{bmatrix}3\\-4\end{bmatrix}$ has

$$\cos\theta = \frac{\begin{bmatrix}1\\2\end{bmatrix}\cdot\begin{bmatrix}3\\-4\end{bmatrix}}{\left\|\begin{bmatrix}1\\2\end{bmatrix}\right\|\left\|\begin{bmatrix}3\\-4\end{bmatrix}\right\|} = \frac{3-8}{\sqrt{5}\sqrt{25}} = -\frac{1}{\sqrt{5}},$$

so $\theta = \cos^{-1}(-1/\sqrt{5}) \approx 2.03\,\text{radians} \approx 116.57^\circ$.

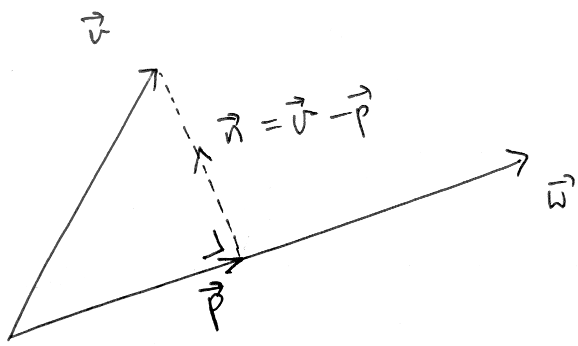
2. The points $A=(2,3)$, $B=(3,6)$ and $C=(-4,5)$ are the vertices of a right-angled triangle. Indeed, we have $\vec{AB} = \begin{bmatrix}3\\6\end{bmatrix} - \begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}1\\3\end{bmatrix}$ and $\vec{AC} = \begin{bmatrix}-4\\5\end{bmatrix} - \begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}-6\\2\end{bmatrix}$, so $\vec{AB}\cdot\vec{AC} = \begin{bmatrix}1\\3\end{bmatrix}\cdot\begin{bmatrix}-6\\2\end{bmatrix} = 1(-6)+3(2)=0$, so the sides AB and AC are at right-angles.

3. To find a unit vector orthogonal to the vector $\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, we can first observe that $\vec{w} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ has $\vec{v} \cdot \vec{w} = 0$, so $\vec{v}$ and $\vec{w}$ are orthogonal; and then consider the vector $\vec{u} = \frac{1}{\|\vec{w}\|}\vec{w}$, which is a unit vector in the same direction as $\vec{w}$, so is also orthogonal to $\vec{v}$. Hence $\vec{u} = \frac{1}{\sqrt{5}}\begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$ is a unit vector orthogonal to $\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

The orthogonal projection of one vector onto another

Let $\vec{w}$ be a non-zero vector, and let $\vec{v}$ be any vector. We call a vector $\vec{p}$ the **orthogonal projection of $\vec{v}$ onto $\vec{w}$**, and write $\vec{p} = \mathrm{proj}_{\vec{w}}\vec{v}$, if

1. $\vec{p}$ is in the same direction as $\vec{w}$; and
2. the vector $\vec{n} = \vec{v} - \vec{p}$ joining the end of $\vec{p}$ to the end of $\vec{v}$ is orthogonal to $\vec{w}$.



We can use these properties of $\vec{p}$ to find a formula for $\vec{p}$ in terms of $\vec{v}$ and $\vec{w}$.

1. Since $\vec{p}$ is in the same direction as $\vec{w}$, we have $\vec{p} = c\vec{w}$ for some scalar $c \in \mathbb{R}$.
2. Since $\vec{n} = \vec{v} - \vec{p}$ is orthogonal to $\vec{w}$, we have $\vec{n} \cdot \vec{w} = 0$. Hence

$$\begin{aligned} & (\vec{v} - \vec{p}) \cdot \vec{w} = 0 \\ \implies \quad & \vec{v} \cdot \vec{w} - \vec{p} \cdot \vec{w} = 0 \\ \implies \quad & \vec{p} \cdot \vec{w} = \vec{v} \cdot \vec{w} \\ \implies \quad & c\vec{w} \cdot \vec{w} = \vec{v} \cdot \vec{w} \\ \implies \quad & c\|\vec{w}\|^2 = \vec{v} \cdot \vec{w} \\ \implies \quad & c = \frac{\vec{v} \cdot \vec{w}}{\|\vec{w}\|^2}. \end{aligned}$$

So

$$\vec{p} = \mathrm{proj}_{\vec{w}}\vec{v} = \frac{\vec{v} \cdot \vec{w}}{\|\vec{w}\|^2}\,\vec{w}.$$

We call $\vec{n} = \vec{v} - \mathrm{proj}_{\vec{w}}\vec{v}$ the component of $\vec{v}$ orthogonal to $\vec{w}$.

Example

If $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$, then

$$\begin{aligned} \mathrm{proj}_{\vec{w}}\vec{v} &= \frac{\vec{v}\cdot\vec{w}}{\|\vec{w}\|^2}\,\vec{w} \\ &= \frac{2-2-4}{2^2+(-1)^2+4^2}\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} \\ &= -\frac{4}{21}\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} \end{aligned}$$

and the component of $\vec{v}$ orthogonal to $\vec{w}$ is

$$\begin{aligned} \vec{n} &= \vec{v} - \mathrm{proj}_{\vec{w}}\vec{v} \\ &= \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} - \left(-\frac{4}{21}\right)\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 8/21 \\ -4/21 \\ 16/21 \end{bmatrix} \\ &= \begin{bmatrix} 29/21 \\ 38/21 \\ -5/21 \end{bmatrix}. \end{aligned}$$

The cross product of vectors in $\mathbb{R}^3$

Definition: the standard basis vectors

We define $\vec{\imath} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{\jmath} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\vec{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. These are the **standard basis vectors** of $\mathbb{R}^3$.

Note that any vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ may be written as a **linear combination** of these vectors (that is, a sum of scalar multiplies of $\vec{\imath}$, $\vec{\jmath}$ and $\vec{k}$), since

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ v_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ v_3 \end{bmatrix} = v_1\vec{\imath} + v_2\vec{\jmath} + v_3\vec{k}.$$

Definition: the cross product

If $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ are vectors in $\mathbb{R}^3$, then we define $\vec{v} \times \vec{w}$ to be the vector given by the determinant

$$\vec{v} \times \vec{w} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

We interpret this determinant by expanding along the first row:

$$\vec{v} \times \vec{w} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} \vec{\imath} - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} \vec{\jmath} + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} \vec{k} = \begin{bmatrix} v_2 w_3 - v_3 w_2 \\ -(v_1 w_3 - v_3 w_1) \\ v_1 w_2 - v_2 w_1 \end{bmatrix}$$

Example

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

$$\vec{v} \times \vec{w} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 3 & -1 \\ 2 & 1 & -2 \end{vmatrix} = \begin{bmatrix} 3(-2) - 1(-1) \\ -(1(-2) - (-1)2) \\ 1(1) - 3(2) \end{bmatrix} = \begin{bmatrix} -5 \\ 0 \\ -5 \end{bmatrix}$$

and

$$\vec{w} \times \vec{v} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 2 & 1 & -2 \\ 1 & 3 & -1 \end{vmatrix} = \begin{bmatrix} 1(-1) - (-2)3 \\ -(2(-1) - (-2)1) \\ 2(3) - 1(1) \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 5 \end{bmatrix}.$$

Observe that $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$. Moreover,

$$\vec{v} \times \vec{v} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 3 & -1 \\ 1 & 3 & -1 \end{vmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0}$$

and

$$\vec{w} \times \vec{w} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 2 & 1 & -2 \\ 2 & 1 & -2 \end{vmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0}.$$

Example: cross products of standard basis vectors

We have

$$\vec{\imath} \times \vec{\jmath} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \vec{k},$$

$$\vec{\jmath} \times \vec{k} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \vec{\imath}$$

$$\vec{k} \times \vec{\imath} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \vec{\jmath}$$

Proposition: properties of the cross product

For any vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^3$ and any scalar $c \in \mathbb{R}$, we have:

1. $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$
2. $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$
3. $(c\vec{v}) \times \vec{w} = c(\vec{v} \times \vec{w}) = \vec{v} \times (c\vec{w})$
4. $\vec{v} \times \vec{v} = \vec{0}$
5. $\vec{v} \times \vec{0} = \vec{0}$
6. $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$

Proof

1. This is a tedious (but easy) bit of algebra.
2. Swapping two rows in a determinant changes the sign, so

$$\vec{v} \times \vec{w} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = -\begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ w_1 & w_2 & w_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = -\vec{w} \times \vec{v}.$$

3. Scaling one row in a determinant scales the determinant in the same way, so

$$(c\vec{v}) \times \vec{w} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ cv_1 & cv_2 & cv_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = c\begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = c\vec{v} \times \vec{w}.$$

4. The determinant of a matrix with a repeated row is zero.
5. The determinant of a matrix with a zero row is zero.
6. Observe that $\vec{u} \cdot (\vec{v} \times \vec{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$. The determinant of a matrix with a repeated row is zero, so

$$\vec{v} \cdot (\vec{v} \times \vec{w}) = \begin{vmatrix} v_1 & v_2 & v_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = 0$$

so $\vec{v}$ is orthogonal to $\vec{v} \times \vec{w}$; and similarly,

$$\vec{w} \cdot (\vec{v} \times \vec{w}) = \begin{vmatrix} w_1 & w_2 & w_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = 0$$

so $\vec{w}$ is orthogonal to $\vec{v} \times \vec{w}$. ■

Theorem

For any vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^3$, we have

$$\|\vec{v} \times \vec{w}\|^2 + (\vec{v} \cdot \vec{w})^2 = \|\vec{v}\|^2 \, \|\vec{w}\|^2.$$

The proof is a tedious but elementary calculation, which we leave as an exercise.

Corollary: the length of $\vec{v} \times \vec{w}$

For any vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^3$, we have

$$\|\vec{v} \times \vec{w}\| = \|\vec{v}\| \, \|\vec{w}\| \, \sin\theta$$

where θ is the angle between $\vec{v}$ and $\vec{w}$ (with $0 \leq \theta < \pi$).

Proof

Recall that $\vec{v} \cdot \vec{w} = \|\vec{v}\| \, \|\vec{w}\| \cos\theta$. Now

$$\begin{aligned} \|\vec{v} \times \vec{w}\|^2 &= \|\vec{v}\|^2 \, \|\vec{w}\|^2 - (\vec{v} \cdot \vec{w})^2 \\ &= \|\vec{v}\|^2 \, \|\vec{w}\|^2 - \|\vec{v}\|^2 \|\vec{w}\|^2 \cos^2\theta \\ &= \|\vec{v}\|^2 \, \|\vec{w}\|^2 (1 - \cos^2\theta) \\ &= \|\vec{v}\|^2 \, \|\vec{w}\|^2 \sin^2\theta. \end{aligned}$$

Since $\sqrt{a^2} = a$ if $a \geq 0$ and $\sin\theta \geq 0$ for $0 \leq \theta < \pi$, taking square roots of both sides gives

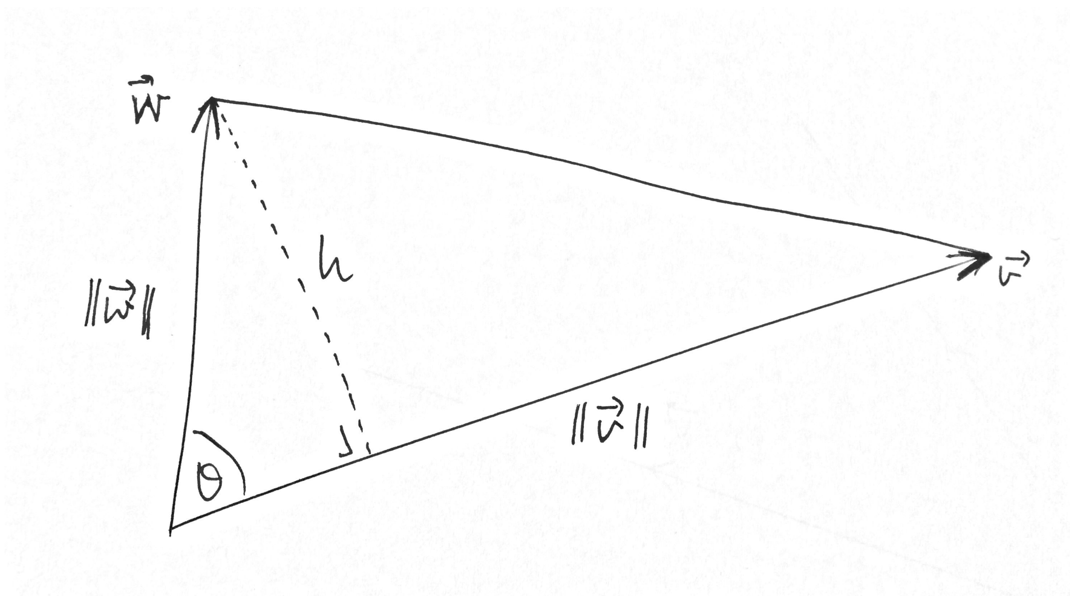
$$\|\vec{v} \times \vec{w}\| = \|\vec{v}\| \, \|\vec{w}\| \, \sin\theta. \blacksquare$$

Geometry of the cross product

Let $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{R}^3$.

The area of a triangle

Consider a triangle with sides $\vec{v}$ and $\vec{w}$ (and a third vector, namely $\vec{v} - \vec{w}$). Thinking of $\vec{v}$ as the base, the length of the base is $b = \|\vec{v}\|$ and the height of this triangle (measured at right angles to the base) is $h = \|\vec{w}\| \sin\theta$ where θ is the angle between $\vec{v}$ and $\vec{w}$.

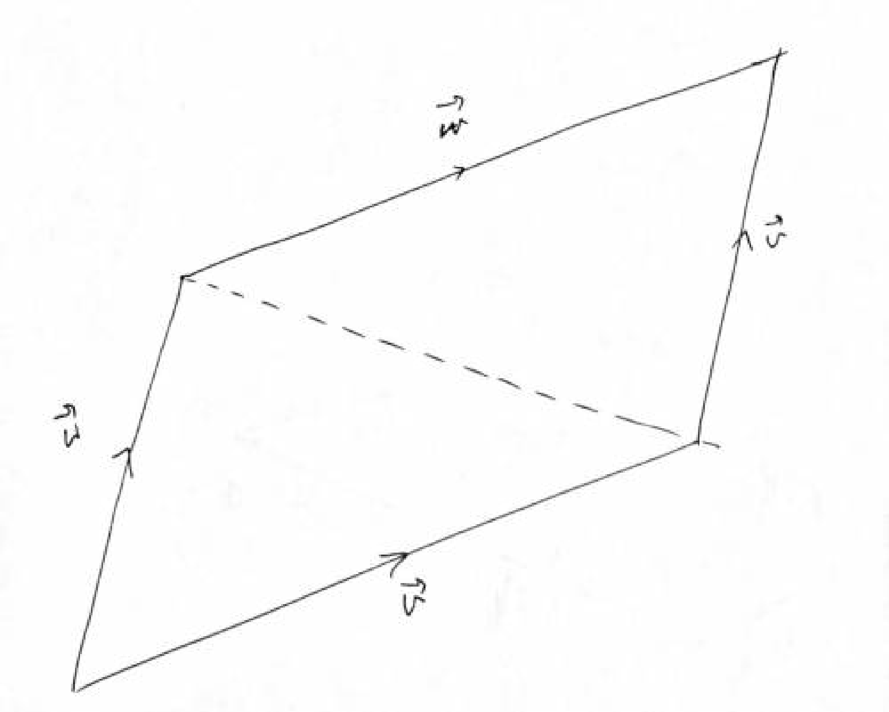


Hence the area of this triangle is $\frac{1}{2}\, bh = \frac{1}{2}\, \|\vec{v}\|\, \|\vec{w}\| \sin\theta$, which is equal to $\frac{1}{2}\, \|\vec{v} \times \vec{w}\|$ (by the formula for $\|\vec{v} \times \vec{w}\|$ which appears above).

MST10030 notes wiki

The area of a parallelogram

Consider a parallelogram, two of whose sides are $\vec{v}$ and $\vec{w}$.



This has double the area of the triangle considered above, so its area is $\|\vec{v} \times \vec{w}\|$.

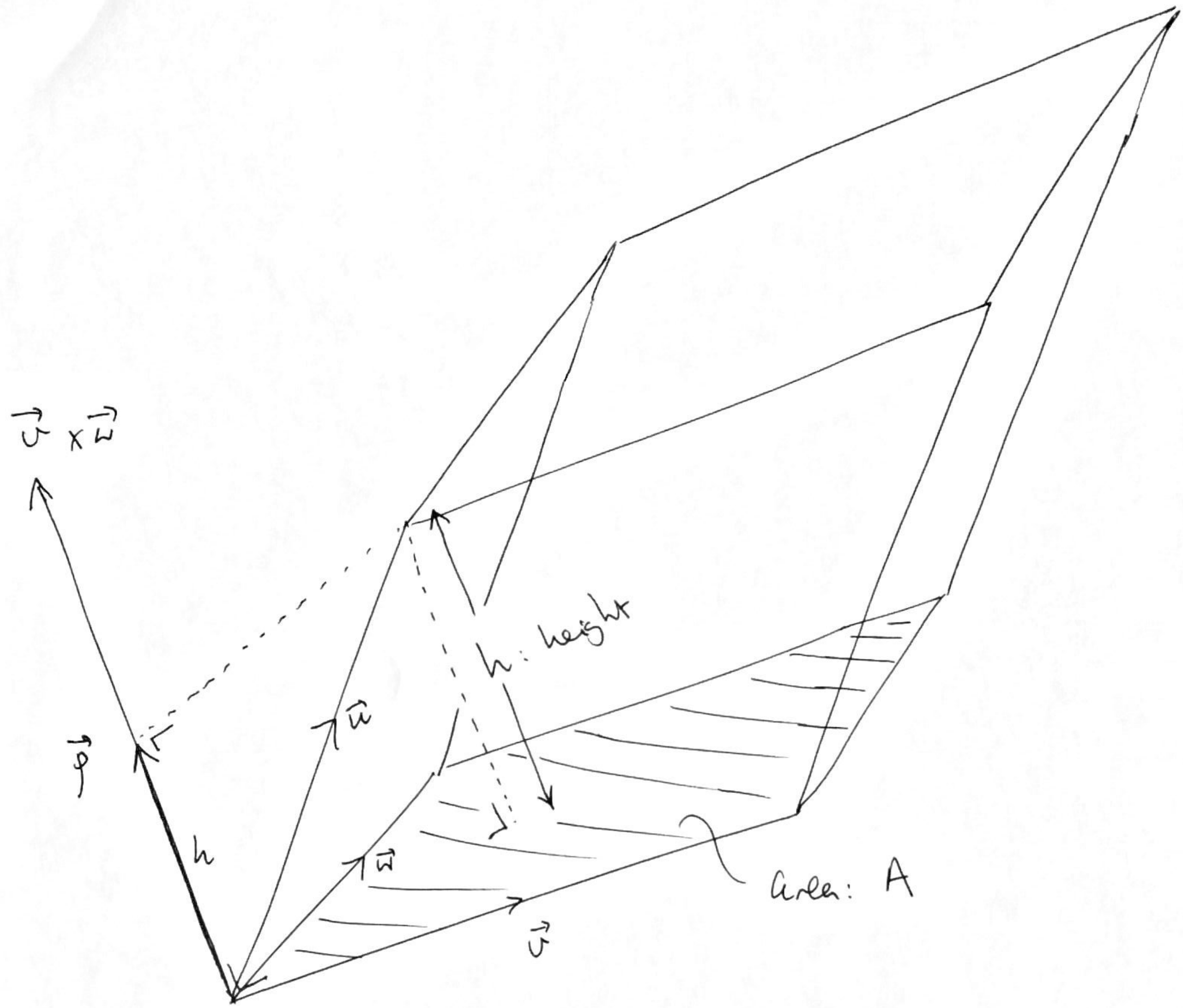
Example

A triangle with two sides $\vec{v} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ has area $\frac{1}{2}\|\vec{v} \times \vec{w}\| = \frac{1}{2}\left\|\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} \times \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}\right\| = \frac{1}{2}\left\|\begin{bmatrix} -5 \\ 0 \\ -5 \end{bmatrix}\right\| = \frac{5}{2}\left\|\begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}\right\| = \frac{5}{2}\sqrt{2}$, and the parallelogram with sides $\vec{v}$ and $\vec{w}$ has area $\|\vec{v} \times \vec{w}\| = 5\sqrt{2}$.

The volume of a parallelepiped in $\mathbb{R}^3$

Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{R}^3$.

Consider a parallelepiped [https://en.wikipedia.org/wiki/parallelepiped], with three sides given by $\vec{u}$, $\vec{v}$ and $\vec{w}$.



Call the face with sides $\vec{v}$ and $\vec{w}$ the base of the parallelpiped. The area of the base is $A = \|\vec{v} \times \vec{w}\|$, and the volume of the parallelpiped is Ah where h is the height, measured at right-angles to the base.

One vector which is at right-angles to the base is $\vec{v} \times \vec{w}$. It follows that h is the length of $\vec{p} = \mathbf{proj}_{\vec{v}\times\vec{w}}\vec{u}$, so

$$h = \|\mathbf{proj}_{\vec{v}\times\vec{w}}\vec{u}\| = \left\| \frac{\vec{u}\cdot(\vec{v}\times\vec{w})}{\|\vec{v}\times\vec{w}\|^2}\,\vec{v}\times\vec{w} \right\| = \frac{\vec{u}\cdot(\vec{v}\times\vec{w})}{\|\vec{v}\times\vec{w}\|^2}\,\|\vec{v}\times\vec{w}\| = \frac{|\vec{u}\cdot(\vec{v}\times\vec{w})|}{\|\vec{v}\times\vec{w}\|}$$

so the volume is

$$V = Ah = \|\vec{v}\times\vec{w}\|\,\frac{|\vec{u}\cdot(\vec{v}\times\vec{w})|}{\|\vec{v}\times\vec{w}\|}$$

or

$$V = |\vec{u}\cdot(\vec{v}\times\vec{w})|.$$

Now $\vec{u}\cdot(\vec{v}\times\vec{w}) = \mathbf{det}\left(\begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix}\right)$, so V is the absolute value of this determinant:

$$V = \left| \quad \mathbf{det}\left(\begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix}\right) \quad \right|.$$

Example

Find volume of a parallelepiped whose vertices include $A=(1,1,1)$, $B=(2,1,3)$, $C=(0,2,2)$ and $D=(3,4,1)$, where A is an adjacent vertex to B, C and D.

Solution

The vectors $\vec{AB}=\begin{bmatrix}1\\0\\2\end{bmatrix}$, $\vec{AC}=\begin{bmatrix}-1\\1\\1\end{bmatrix}$ and $\vec{AD}=\begin{bmatrix}2\\3\\0\end{bmatrix}$ are all edges of this parallepiped, so the volume is

$$V=\left|\begin{vmatrix}1&0&2\\-1&1&1\\2&3&0\end{vmatrix}\right|=|1(0-3)-0+2(-3-2)|=|-13|=13.$$

Planes and lines in $\mathbb{R}^3$

Recall that a typical plane in $\mathbb{R}^3$ has equation

$$ax+by+cz=d$$

where a,b,c,d are constants. If we write

$$\vec{n}=\begin{bmatrix}a\\b\\c\end{bmatrix}$$

then we can rewrite the equation of this plane in the form

$$\vec{n}\cdot\begin{bmatrix}x\\y\\z\end{bmatrix}=d.$$

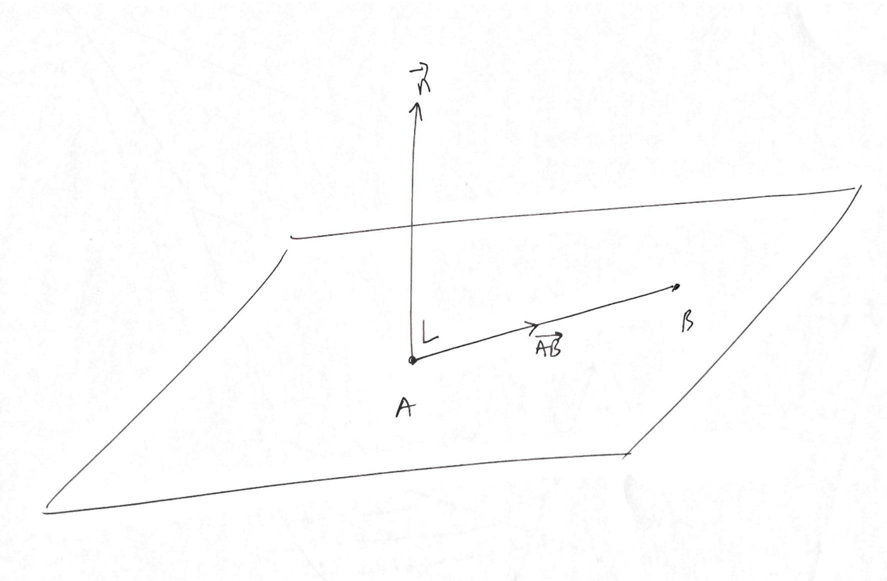
If $A=(x_1,y_1,z_1)$ and $B=(x_2,y_2,z_2)$ are both points in this plane, then the vector $\vec{AB}$ is said to be **in the plane**, or to be **parallel to** the plane. Observe that

$$\vec{n}\cdot\vec{AB}=\vec{n}\cdot\left(\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}-\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}\right)=\vec{n}\cdot\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}-\vec{n}\cdot\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}=d-d=0,$$

so

$$\vec{n}\cdot\vec{v}=0$$

for every vector $\vec{v}$ in the plane. In other words: the vector $\vec{n}$ is orthogonal to every vector in the plane.



We call a vector with this property a **normal** vector to the plane.

Examples

1. Find a unit normal vector to the plane $x+y-3z=4$.

Solution: The vector $\vec{n}=\begin{bmatrix}1\\1\\-3\end{bmatrix}$ is a normal vector to this plane, so $\vec{v}=\frac{1}{\|\vec{n}\|}\vec{n}=\frac{1}{\sqrt{11}}\begin{bmatrix}1\\1\\-3\end{bmatrix}$ is a unit normal vector to this plane. Indeed, $\vec{v}$ is a unit vector and it's in the same direction as the normal vector $\vec{n}$, so $\vec{v}$ is also a normal vector.

2. Find the equation of the plane with normal vector $\begin{bmatrix}1\\-3\\2\end{bmatrix}$ which contains the point $(1,-2,1)$. Then find three other points in this plane.

Solution: the equation is $x-3y+2z=d$, and we can find d by subbing in $(x,y,z)=(1,-2,1)$: $1-3(-2)+2(1)=d$, so $d=9$ and the equation of the plane is

$$x-3y+2z=9.$$

Some other points in this plane are $(9,0,0)$, $(0,1,6)$, $(1,1,\frac{11}{2})$. (We can find these by inspection).

3. Find the equation of the plane parallel to the vectors $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\-1\\1\end{bmatrix}$ containing the point $(3,0,1)$.

Solution: a normal vector is $\vec{n}=\begin{bmatrix}1\\1\\1\end{bmatrix}\times\begin{bmatrix}1\\-1\\1\end{bmatrix}=\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}111-11=\begin{bmatrix}2\\0\\-2\end{bmatrix}$, so the equation is $2x+0y-2z=2(3)-2(1)=4$, or $2x-2z=4$, or $x-z=2$.

4. Find the equation of the plane containing the points $A=(1,2,0)$, $B=(3,0,1)$ and $C=(4,3,-2)$.

Solution: $\vec{AB}=\begin{bmatrix}2\\-2\\1\end{bmatrix}$ and $\vec{AC}=\begin{bmatrix}3\\1\\-2\end{bmatrix}$ are both vectors in this plane. We want to find a normal vector $\vec{n}$ which is be orthogonal to both of these. The cross product of two vectors is orthogonal to both, so we can take the cross product of $\vec{AB}$ and $\vec{AC}$:

$$\vec{n}=\vec{AB}\times\vec{AC}=\begin{vmatrix}\vec{i}&\vec{j}&\vec{k}\\2&-2&1\\3&1&-2\end{vmatrix}=\begin{bmatrix}3\\7\\8\end{bmatrix}$$

so the equation of the plane is $3x+7y+8z=d$, and we find d by subbing in a point in the plane, say $A=(1,2,0)$, which gives $d=3(1)+7(2)+8(0)=17$. So the equation is

$$3x+7y+8z=17.$$

Orthogonal planes and parallel planes

Let Π_1 be a plane with normal vector $\vec{n}_1$, and let Π_2 be a plane with normal vector $\vec{n}_2$.

1. Π_1 and Π_2 are *orthogonal* or *perpendicular* planes if they meet at right angles. The following conditions are equivalent:
   a. Π_1 and Π_2 are orthogonal planes;
   b. $\vec{n}_1\cdot\vec{n}_2=0$;
   c. $\vec{n}_1$ is a vector in Π_2;
   d. $\vec{n}_2$ is a vector in Π_1.
2. Π_1 and Π_2 are *parallel* planes if they have the same normal vectors. In other words, if Π_1 has equation $ax+by+cz=d_1$ then any parallel plane Π_2 has an equation with the same left hand side: $ax+by+cz=d_2$.

Examples

1. Find the equation of the plane Π passing through $A=(1,3,-3)$ and $B=(4,-2,1)$ which is orthogonal to the plane $x-y+z=5$.

Solution: The plane $x-y+z=5$ has normal vector $\begin{bmatrix}1\\-1\\1\end{bmatrix}$, so this is a vector in Π. Moreover, $\vec{AB}=\begin{bmatrix}3\\-5\\4\end{bmatrix}$ is also a vector in Π, so it has normal vector

$$\vec{n}=\begin{bmatrix}1\\-1\\1\end{bmatrix}\times\begin{bmatrix}3\\-5\\4\end{bmatrix}=\begin{vmatrix}\vec{i}&\vec{j}&\vec{k}\\1&-1&1\\3&-5&4\end{vmatrix}=\begin{bmatrix}1\\-1\\-2\end{bmatrix}.$$

So the equation of Π is $x-y-2z=d$ and subbing in $A=(1,3,-3)$ gives $d=1-3-2(-3)=4$, so the equation of Π is

$$x - y - 2z = 4.$$

2. The plane parallel to $2x - 4y + 5z = 8$ passing through $(1, 2, 3)$ is
$2x - 4y + 5z = 2(1) - 4(2) + 5(3) = 10$, or $2x - 4y + 5z = 10$.

3. Find the equation of the plane Π which contains the line of intersection of the planes

$$\Pi_1 : x - y + 2z = 1 \quad \text{and} \quad \Pi_2 : 3x + 2y - z = 4,$$

and is perpendicular to the plane $\Pi_3 : 2x + y + z = 3$.

Solution: To find the line of intersection of Π_1 and Π_2, we must solve the system of linear equations

$$\begin{aligned} x - y + 2z &= 1 \\ 3x + 2y - z &= 4. \end{aligned}$$

We can solve this linear system in the usual way, by applying EROs to the matrix $\begin{bmatrix} 1 & -1 & 2 & 1 \\ 3 & 2 & -1 & 4 \end{bmatrix}$:

$$\begin{aligned} & \begin{bmatrix} 1 & -1 & 2 & 1 \\ 3 & 2 & -1 & 4 \end{bmatrix} \\ \xrightarrow{R2 \to R2 - 3R1} & \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 5 & -7 & 1 \end{bmatrix} \\ \xrightarrow{R1 \to 5R1 + R2} & \begin{bmatrix} 5 & 0 & 3 & 6 \\ 0 & 5 & -7 & 1 \end{bmatrix} \\ \xrightarrow{R1 \to \frac{1}{5} R1,\, R2 \to \frac{1}{5} R2} & \begin{bmatrix} 1 & 0 & 3/5 & 6/5 \\ 0 & 1 & -7/5 & 1/5 \end{bmatrix} \end{aligned}$$

So the line L of intersection is given by

$$L : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{6}{5} \\ \frac{1}{5} \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{5} \\ \frac{7}{5} \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

So $\begin{bmatrix} -\frac{3}{5} \\ \frac{7}{5} \\ 1 \end{bmatrix}$ is a direction vector along L, and also $5 \begin{bmatrix} -\frac{3}{5} \\ \frac{7}{5} \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 7 \\ 5 \end{bmatrix}$ is a vector along L. So $\begin{bmatrix} -3 \\ 7 \\ 5 \end{bmatrix}$ is a vector in the plane Π. Moreover, taking $t = 2$ gives the point $(0, 3, 2)$ in the line L, so this is a point in Π.

Since Π is perpendicular to Π_3, which has normal vector $\vec{n}_3 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, the vector $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ is in Π.

So a normal vector for Π is

$$\vec{n} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} -3 \\ 7 \\ 5 \end{bmatrix} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 2 & 1 & 1 \\ -3 & 7 & 5 \end{vmatrix} = \begin{bmatrix} -2 \\ -13 \\ 17 \end{bmatrix}$$

hence Π has equation $-2x - 13y + 17z = d$, and subbing in the point $(0, 3, 2)$ gives $d = 0 - 13(3) + 17(2) = -39 + 34 = -5$, so Π has equation $-2x - 13y + 17z = -5$, or

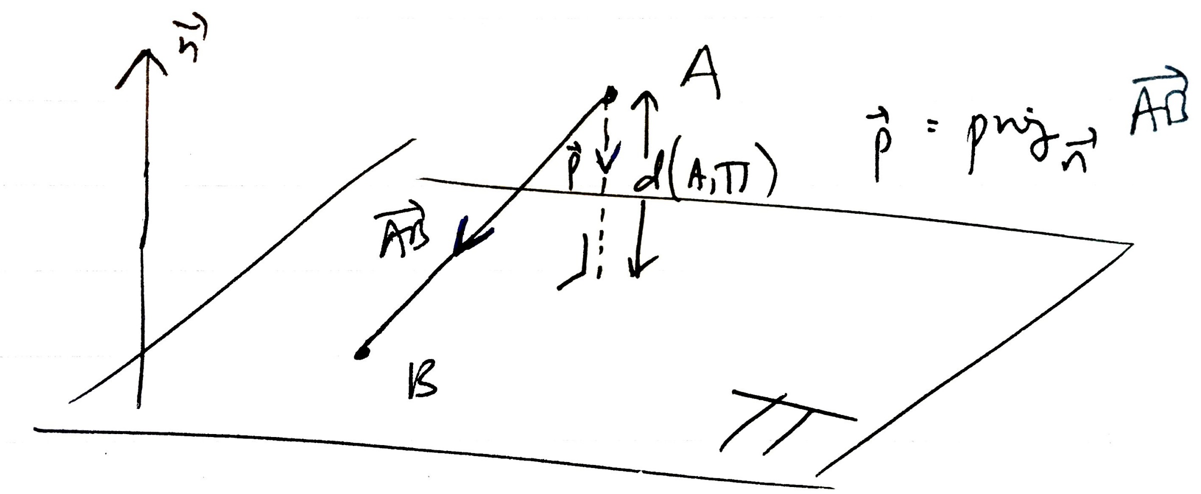
$$2x + 13y - 17z = 5.$$

The distance to a plane

The distance from a point to a plane

Let Π be a plane in $\mathbb{R}^3$ with equation $ax + by + cz = d$, so that $\vec{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is a normal vector to Π. Also let A be any point in $\mathbb{R}^3$.

The shortest path from A to a point in Π goes in the same direction as $\vec{n}$. Let B be any point in the plane Π.



From the diagram, we see that the shortest distance from A to Π is given by

$$\text{dist}(A, \Pi) = \|\vec{p}\|$$

where

$$\vec{p} = \text{proj}_{\vec{n}} \vec{AB}.$$

Using the formula for $\text{proj}_{\vec{w}} \vec{v}$ and the fact that $\|c\vec{v}\| = |c| \, \|\vec{v}\|$ where c is a scalar and $\vec{v}$ is a vector, we obtain the formula

$$\text{dist}(A, \Pi) = \frac{|\vec{n} \cdot \vec{AB}|}{\|\vec{n}\|}.$$

Example

To find the distance from $A = (1, -4, 3)$ to the plane $\Pi : 2x - 3y + 6z = 1$, choose any point B in Π; for example, let $B = (2, 1, 0)$. Then $\vec{n} = \begin{bmatrix} 2 \\ -3 \\ 6 \end{bmatrix}$ and $\vec{AB} = \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}$, so

$$\operatorname{dist}(A,\Pi)=\frac{|\vec n\cdot \vec{AB}|}{\|\vec n\|}=\frac{|2(1)+(-3)5+6(-3)|}{\sqrt{2^2+(-3)^2+6^2}}=\frac{|-31|}{\sqrt{49}}=\frac{31}{7}\,.$$

Remark: the distance from the origin to a plane

If we write $0 = (0, 0, 0)$ for the origin in $\mathbb{R}^3$ and apply the formula above to the plane $\Pi : ax + by + cz = d$ with $B = (d/a, 0, 0)$ (assuming that $a \neq 0$) then we obtain

$$\text{dist}(0, \Pi) = \frac{|d|}{\|\vec{n}\|}$$

where $\vec{n}$ is the normal vector $\vec{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

So as d varies (with the normal vector $\vec{n}$ fixed), we obtain parallel planes at different distances to the origin 0; the larger d is, the further the plane is from 0.

The distance between parallel planes

If Π_1 and Π_2 are parallel planes, then the shortest distance between them is given by

$$\text{dist}(\Pi_1, \Pi_2) = \text{dist}(A, \Pi_2)$$

for any point A is Π_1. The reason is that for parallel planes, changing A to a different point in Π_1 does not change $\text{dist}(A, \Pi_2)$.

Of course, if the planes Π_1 and Π_2 are not parallel, then they intersect (in many points: in a whole line). So for non-parallel planes we always have $\text{dist}(\Pi_1, \Pi_2) = 0$.

Example

The distance between the planes $3x + 4y - 2z = 5$ and $3x + 4y - 3z = 1$ is 0, since the normal vectors $\begin{bmatrix} 3 \\ 4 \\ -2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 4 \\ -3 \end{bmatrix}$ are not scalar multiples of one another, so they are in different directions, so the planes are not parallel.

Example

The planes $\Pi_1 : 3x + 4y - 2z = 5$ and $\Pi_2 : 3x + 4y - 2z = 1$ have the same normal vector $\begin{bmatrix} 3 \\ 4 \\ -2 \end{bmatrix}$, so they are parallel. Their distance is given by $\text{dist}(A, \Pi_2)$ where A is any point in Π_1, and to find this we also need a point B in Π_2.

We can choose $A = (1, 0, -1) \in \Pi_1$ and $B = (1, 0, 1) \in \Pi_2$. (Of course, there are lots of different possible choices here, but they should all give the same answer!) Then $\vec{AB} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$ and

$$\text{dist}(\Pi_1, \Pi_2) = \text{dist}(A, \Pi_2) = \frac{|\vec{n} \cdot \vec{AB}|}{\|n\|} = \frac{|0 + 0 + (-2)2|}{\sqrt{3^2 + 4^2 + (-2)^2}} = \frac{4}{\sqrt{29}}\,.$$

Exercise: a formula for the distance between parallel planes

Show that the distance between the parallel planes $\Pi_1 : ax + by + cz = d_1$ and $\Pi_2 : ax + by + cz = d_2$ is

$$\text{dist}(\Pi_1, \Pi_2) = \frac{|d_2 - d_1|}{\|\vec{n}\|}\,,$$

where $\vec{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

Example

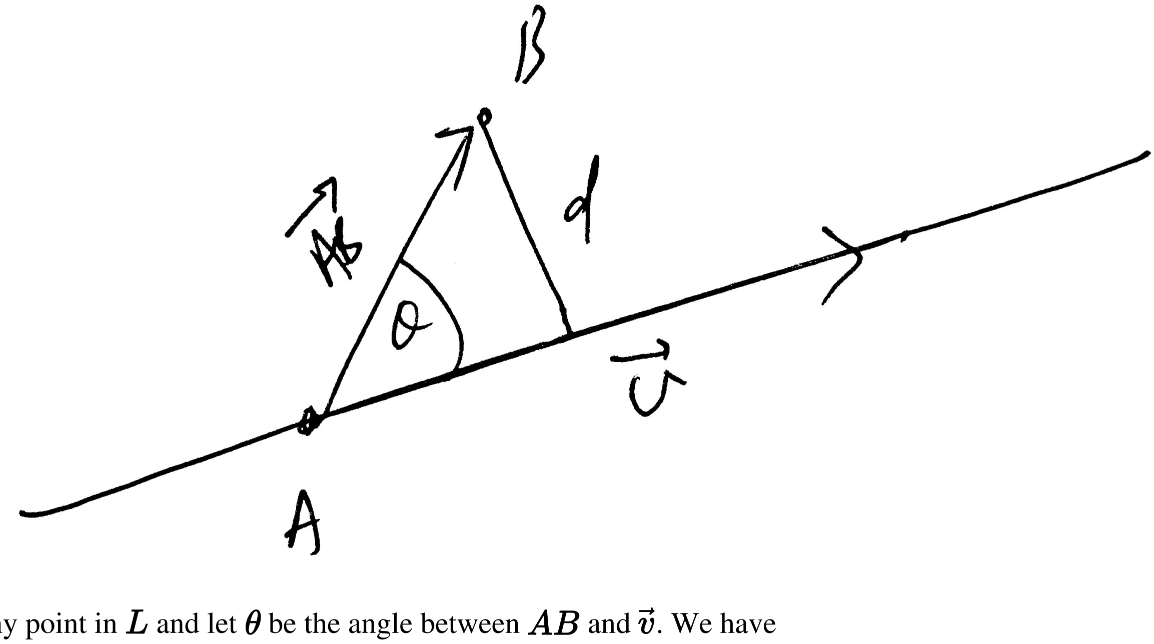
To find the distance between $x + 3y - 5z = 4$ and $2x + 6y - 10z = 11$ we can rewrite the second equation as $x + 3y - 5z = 11/2$ to see that this is a parallel plane to the first, with common normal vector $\vec{n} = \begin{bmatrix} 1 \\ 3 \\ -5 \end{bmatrix}$. By the formula in the exercise the distance between these planes is

$$\frac{|\,\frac{11}{2} - 4|}{\|\vec{n}\|} = \frac{|\,\frac{3}{2}\,|}{\sqrt{1^2 + 3^2 + (-5)^2}} = \frac{3}{2\sqrt{35}}\,.$$

The distance from a point to a line

Suppose L is a line in $\mathbb{R}^3$. Let A be a point on L and let $\vec{v}$ be a direction vector along L.

Given a point B, how can we find $d = \text{dist}(B, L)$, the (shortest) distance from the point B to the line L?



Let A be any point in L and let θ be the angle between AB and $\vec{v}$. We have

$$d = \|\vec{AB}\|\ \sin\theta = \frac{\|\vec{AB}\|\ \|\vec{v}\|\sin\theta}{\|\vec{v}\|} = \frac{\|\vec{AB}\times\vec{v}\|}{\|\vec{v}\|}\ .$$

So

$$\operatorname{dist}(B,L) = \frac{\|\vec{AB}\times\vec{v}\|}{\|\vec{v}\|}$$

where A is any point in L.

Example

To find the distance from the point $B=(1,2,3)$ to the line

$$L: \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + t\begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix}, \quad t\in\mathbb{R}$$

we can choose $A = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ so that $\vec{AB} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}$ and taking $\vec{v} = \begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix}$, we obtain

$$\vec{AB}\times\vec{v} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 0 & 2 & 4 \\ 4 & 1 & -5 \end{vmatrix} = \begin{bmatrix} -14 \\ 16 \\ -8 \end{bmatrix} = 2\begin{bmatrix} -7 \\ 8 \\ -4 \end{bmatrix}$$

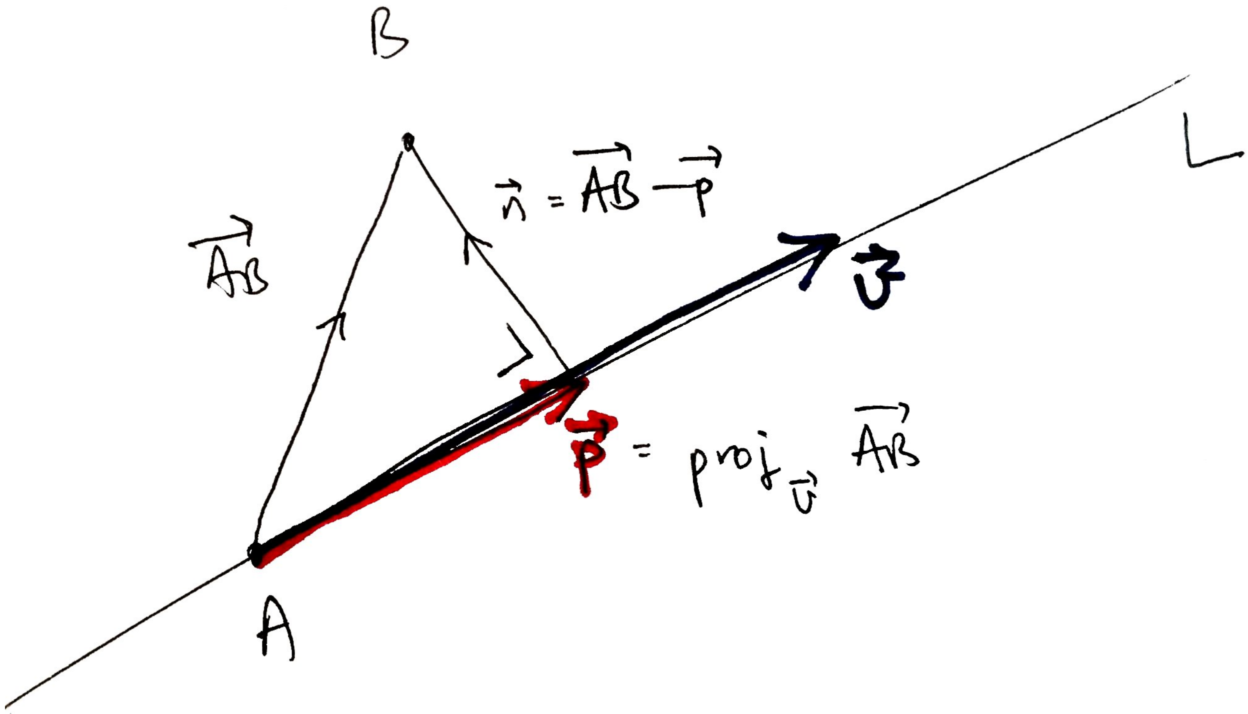
so

$$\operatorname{dist}(B,L) = \frac{\|\vec{AB}\times\vec{v}\|}{\|\vec{v}\|} = \frac{2\sqrt{7^2+8^2+4^2}}{\sqrt{4^2+1^2+5^2}} = \frac{2\sqrt{129}}{\sqrt{42}} \approx 3.5051.$$

The distance from a point to a line

Alternative method

The method above relies on the cross product, so only works in $\mathbb{R}^3$. The following alternative method works in $\mathbb{R}^n$ for any n.



Observe that $\text{dist}(B, L)$ is the length of the vector $\vec{n} = \vec{AB} - \vec{p}$ where $\vec{p} = \textbf{proj}_{\vec{v}}\vec{AB}$.

Example

Let's redo the previous example using this method.

We have $\vec{AB} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix}$, so

$$\vec{p} = \textbf{proj}_{\vec{v}}\vec{AB} = \left(\frac{\vec{AB}\cdot\vec{v}}{\|\vec{v}\|^2}\right)\vec{v} = \frac{0(4)+2(1)+4(-5)}{4^2+1^2+5^2}\begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix} = -\frac{18}{42}\begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix} = -\frac{3}{7}\begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix},$$

so

$$\vec{n} = \vec{AB} - \vec{p} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix} - (-\frac{3}{7})\begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix} + \frac{3}{7}\begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 12 \\ 17 \\ 13 \end{bmatrix}$$

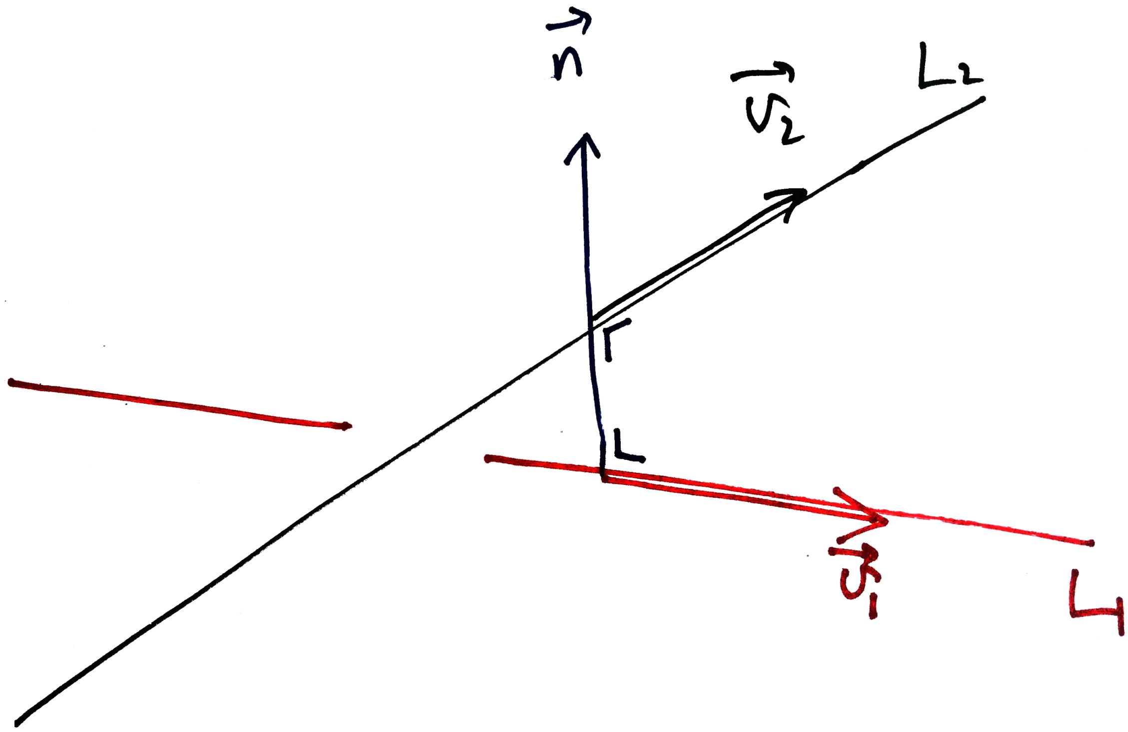
so

$$\text{dist}(B, L) = \|\vec{n}\| = \frac{1}{7}\sqrt{12^2 + 17^2 + 13^2} = \frac{1}{7}\sqrt{602} \approx 3.5051.$$

The distance between skew lines in $\mathbb{R}^3$

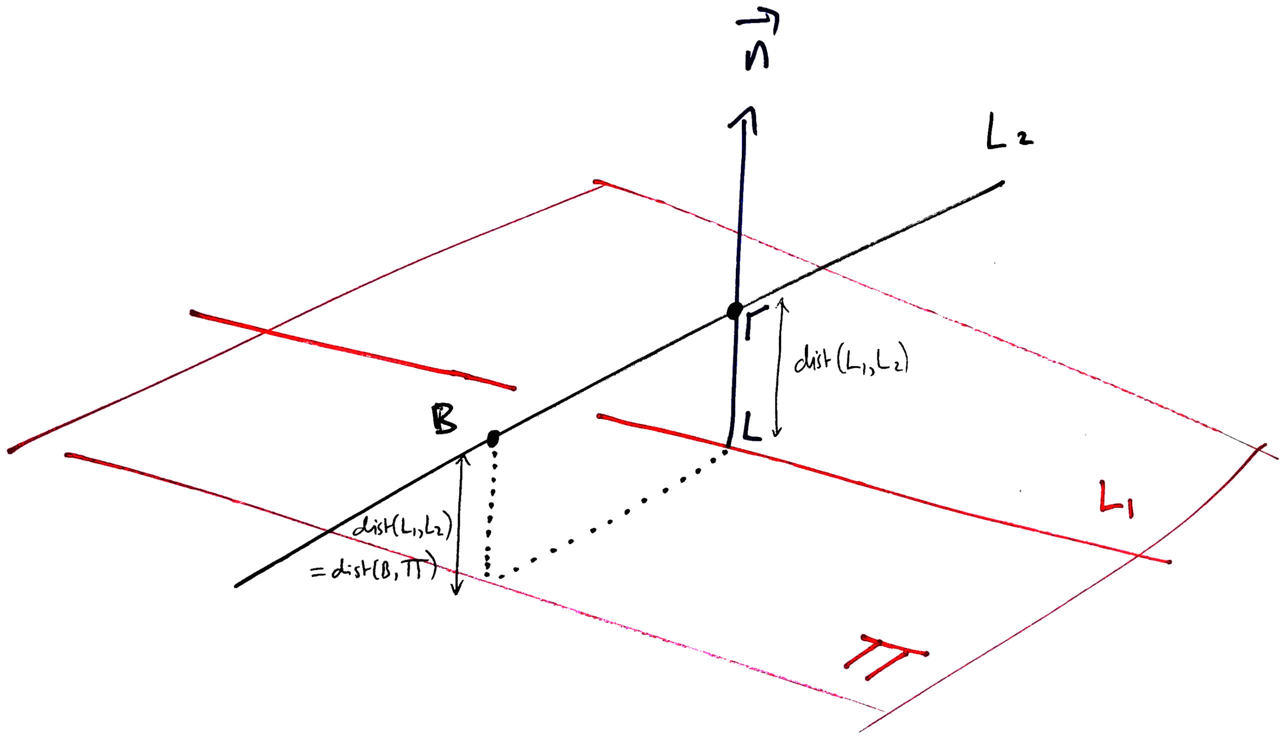
Suppose that L_1 and L_2 are skew lines in $\mathbb{R}^3$: lines which are not parallel and do not cross.

Let $\vec{v}_1$ be a direction vector along L_1, and let $\vec{v}_2$ be a direction vector along L_2.



The shortest distance from L_1 and L_2 is measured along the direction orthogonal to both $\vec{v}_1$ and $\vec{v}_2$, namely the direction of $\vec{n} = \vec{v}_1 \times \vec{v}_2$.

Let Π be the plane with normal vector $\vec{n}$ which contains L_1.



For any point B in L_2, we have

$$\text{dist}(L_1, L_2) = \text{dist}(B, \Pi) = \frac{|\vec{AB} \cdot \vec{n}|}{\|\vec{n}\|}$$

where A is any point in Π; for example, we can take A to be any point in L_1.

To summarise: for skew lines L_1 and L_2 with direction vectors $\vec{v}_1$ and $\vec{v}_2$, we have

$$\text{dist}(L_1, L_2) = \frac{|\vec{AB} \cdot \vec{n}|}{\|\vec{n}\|}$$

where $\vec{n} = \vec{v}_1 \times \vec{v}_2$ and A and B are points with one in L_1 and the other in L_2.

Remark

What about the distance between lines which are not skew? This means that either they are non-parallel and they intersect (so that the distance between them will be zero), or the are parallel lines.

- The same method and formula work if L_1 and L_2 are non-parallel lines which intersect, and you get $\text{dist}(L_1, L_2) = 0$ in this case. The reason is that in this case L_1 and L_2 will lie in one plane, Π, and $\vec{AB}$ will also be in Π, and $\vec{n}$ will be orthogonal to Π. So $\frac{\vec{AB} \cdot \vec{n}}{\|\vec{n}\|} = \frac{0}{\|\vec{n}\|} = 0 = \text{dist}(L_1, L_2)$.
- If L_1 and L_2 are parallel lines (i.e., if the vectors $\vec{v}_1$ and $\vec{v}_2$ along the lines are in the same direction), then $\vec{v}_1 \times \vec{v}_2 = 0$ which isn't helpful, so this method won't work here. In this case, observe that $\text{dist}(L_1, L_2) = \text{dist}(A, L_2)$ where A is any point in L_1 (because the lines are parallel), so you can use one of the formulae above for the distance from a point to a line.

Example

Consider the skew lines

$$L_1: \quad \begin{aligned} x &= 1 + t_1 \\ y &= 2t_1 \qquad (t_1 \in \mathbb{R}) \\ z &= 1 + 3t_1 \end{aligned}$$

and

$$L_2: \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + t_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad t_2 \in \mathbb{R}.$$

Note that we can rewrite the equation of L_1 in "vector form", which is easier to digest:

$$L_1: \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + t_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad t_1 \in \mathbb{R}$$

The direction vectors are $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, so we take $\vec{n}$ to be their cross product:

$$\vec{n} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \times \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 2 & 3 \\ 1 & -1 & 1 \end{vmatrix} = \begin{bmatrix} 5 \\ 2 \\ -3 \end{bmatrix}$$

and if $A = (1,0,1)$ and $B = (3,2,1)$ then A and B are points with one in L_1 and the other in L_2, and $\vec{AB} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$. Hence

$$\operatorname{dist}(L_1, L_2) = \frac{|\vec{AB} \cdot \vec{n}|}{\|\vec{n}\|} = \frac{2(5) + 2(2) + 0(-3)}{\sqrt{5^2 + 2^2 + 3^2}} = \frac{14}{\sqrt{38}}.$$