MAT140: Lecture 16 Handout
*Linear Algebra*

---

Last lecture we introduced matrices and determinants. Matrices first appeared as a convenient way to organise the coefficients of a system of linear equations, and determinants then gave us a way to extract a single number from a square matrix. We also saw that determinants were not merely some "algebraic formula to remember", but instead they described geometric quantities such as area.

In this lecture we will complete our discussion of systems of linear equations by taking a step back and looking at the bigger picture of everything we have done. As a matter of fact, throughout our discussion in the past three lectures we have been flirting with the field of *linear algebra*, which studies vectors, vector spaces, matrices, linear transformations, and systems of linear equations. Today we will start to properly introduce this topic, by tying everything we have seen so far together into a linear-algebraic narrative. Generally speaking, we will focus on the geometric narratives underlying everything.

Today we will:

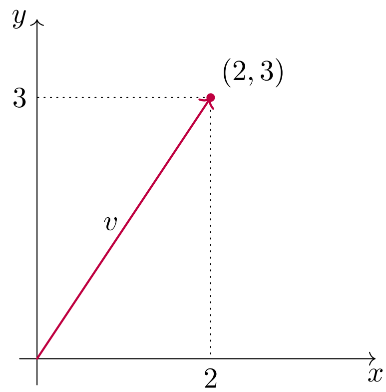
1. Define new objects called *vectors*, and explain how to manipulate them.
2. Describe span and linear combinations, and introduce the standard basis vectors.
3. Define matrix addition and multiplication, and explain why matrix multiplication is unusual.
4. Introduce linear transformations and show how every such transformation is described by a matrix.
5. Reinterpret determinants as area scale factors, and connect matrices back to the solution sets of linear systems.

This lecture should be interpreted as somewhat supplementary. We will not go deeply into these subjects, but instead introduce ideas and connect them to our previous intuitions. For those of you students who are looking to commit to more technical fields of study, you may consider this lecture as a healthy introduction to the ideas you will need to learn in depth later on. For those of you students who are not looking to study deeper mathematics than this, then you may consider this lecture as the mathematical apex of the course, the great mountaintop that we have been climbing towards for the past few lectures. Of course, these perspectives are not mutually exclusive.

# 1 Vectors

## 1.1 What is a vector?

A vector is a mathematical object that has both *magnitude* and *direction*. In the 2-dimensional plane, we often picture a vector as an arrow based at the origin. Given that the tip of this little arrow has to land *somewhere*, we can understand vectors and points as partners of each other. For example, the point $(2, 3)$ determines the vector $v$ which may be drawn as the arrow from $(0, 0)$ to $(2, 3)$:




So, the vector $v$ is the unique arrow starting at the origin and ending at the point $(2,3)$. The two pieces of information "$x=2$" and "$y=3$" uniquely specify the vector $v$, so it makes sense to represent the vector in terms of these two numbers, written in some kind of order. For future convenience, we will write vectors with *vertical* parentheses instead of horizontal ones. In the case of the vector $v$ drawn above, we write:

$$v = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

**The Point-Vector Correspondence**

In two dimensions, there is a natural correspondence between points and vectors: the point $(x,y)$ corresponds to the vector

$$v = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Here $v$ can be interpreted as the directed arrow from the origin to the point $(x,y)$.

A *vector space* is a collection of vectors in which two basic operations make sense:

1. vectors can be added together, and
2. vectors can be multiplied by real numbers, which are called *scalars*.

In this lecture we will mostly focus on the case of 2-dimensional and 3-dimensional vectors that are scaled using real numbers. Mathematically, these two vector spaces are built from $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. We will now detail the nature of these addition and multiplication operations.

## 1.2 Vector arithmetic

If

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \qquad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

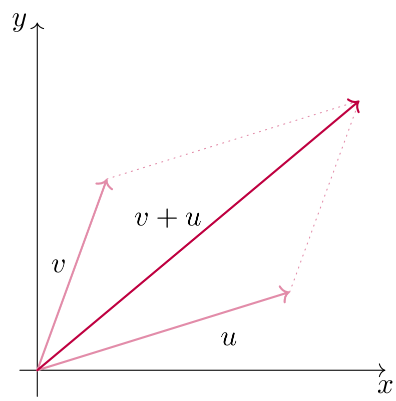
are a pair of vectors, then we can add them componentwise, i.e. by adding their $x$ components to each other, and then adding their $y$-components together:

$$u + v = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}.$$

Likewise, if $a$ is a real number, then scalar multiplication is given by

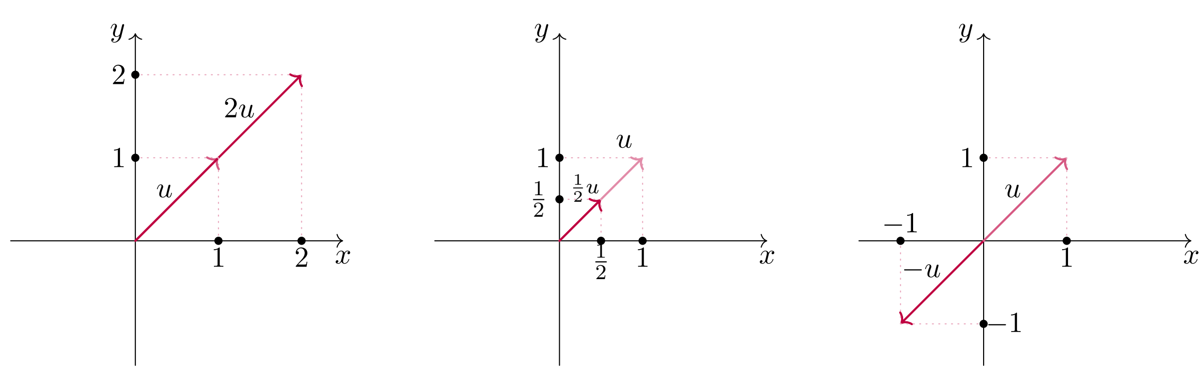
$$au = \begin{pmatrix} au_1 \\ au_2 \end{pmatrix}.$$

Geometrically, vector addition corresponds to placing arrows tip-to-tail. Doing so creates a parallelogram, for which the sum of two vectors is the diagonal arrow from the origin to the newly-created vertex:




Observe in the picture above that placing a copy of $u$ onto the tip of $v$ and placing a copy of $v$ onto the tip of $u$ both lead to the same opposite vertex of the parallelogram. This demonstrates that vector addition is commutative, i.e. $v + u = u + v$.

Scalar multiplication simply stretches, shrinks, or reverses a vector's direction. For example, consider the vector $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. The vectors $2u$, $\frac{1}{2}u$ and $-u$ are drawn below:




As you can see, multiplying by 2 doubles the length, multiplying by $\frac{1}{2}$ halves its length, and multiplying by $-1$ keeps the length the same but reverses the direction of $u$.

We will now present an example of the addition of two vectors. Let

$$u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad v = \begin{pmatrix} -2 \\ 3 \end{pmatrix}.$$

The vectors $u+v$,$v-u$ and $3u-v$ can all be computed by performing these operations componentwise. We have:

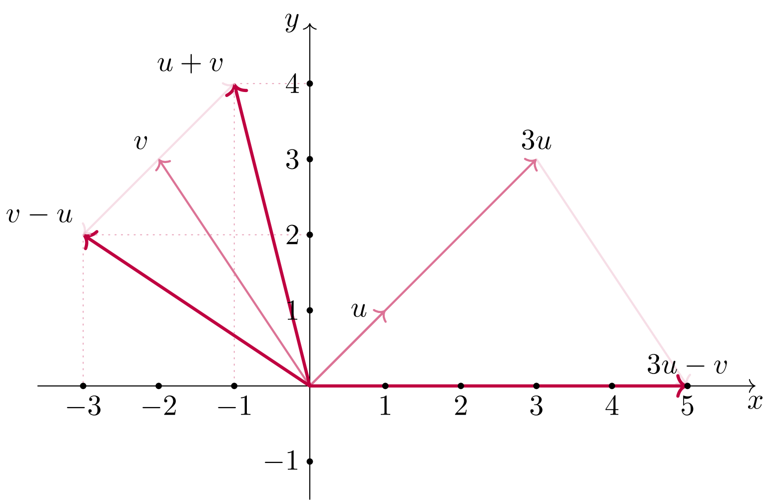
$$u+v = \begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}-2\\3\end{pmatrix} = \begin{pmatrix}-1\\4\end{pmatrix},$$

$$v-u = \begin{pmatrix}-2\\3\end{pmatrix} - \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}-3\\2\end{pmatrix},$$

and

$$3u-v = 3\begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}-2\\3\end{pmatrix} = \begin{pmatrix}3\\3\end{pmatrix} - \begin{pmatrix}-2\\3\end{pmatrix} = \begin{pmatrix}5\\0\end{pmatrix}.$$

Graphically, these vectors look like:




## 1.3 Basic properties of vectors

The rules of vector arithmetic can be collected together into the formal definition of a vector space. We will not emphasise full abstraction here, but the following properties are worth remembering.

**Basic properties of vectors**

For vectors $u, v, w$ and scalars $a, b$, the following hold:

1. There is a zero vector 0.
2. Commutative addition: $u + v = v + u$.
3. Associativity of addition: $(u + v) + w = u + (v + w)$.
4. Additive inverse: $u + (-u) = 0$.
5. Distributivity: $a(u + v) = au + av$.
6. Distributivity: $(a + b)u = au + bu$.
7. Associativity of scalar multiplication: $(ab)u = a(bu)$.
8. Scalar identity: $1u = u$.

Of course, all of these properties may be verified geometrically. For example, we saw previously that adding one vector $v$ to another vector $w$ corresponds to a new vector which is drawn along the diagonal of the parallelogram formed from $v$ and $w$. It does not matter which way round you add the two vectors together, because both tip-to-tail constructions lead to the same opposite vertex of the parallelogram. Algebraically, this can be easily realised by noticing that vector addition simply adds components to components, and therefore it inherits the commutativity property from the real numbers:

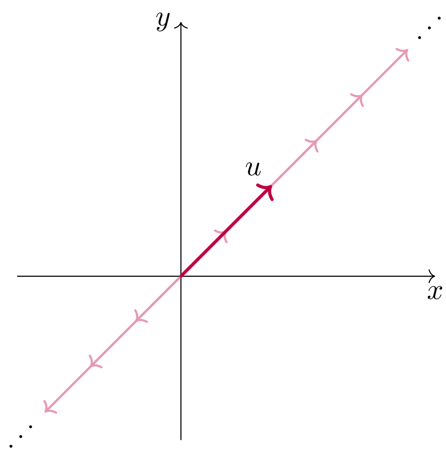
$$u + v = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix} \overset{(*)}{=} \begin{pmatrix} v_1 + u_1 \\ v_2 + u_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = v + u,$$

where here in the step $(*)$ we have simply used that addition of real numbers is commutative.

## 1.4 Span and linear combinations

If $v$ is a nonzero vector, then all scalar multiples of $v$ form a line passing through the origin. This set is called the *span* of $v$:

$$\text{span}(v) = \{\text{all scalar multiples of } v\}.$$




More generally, given two vectors $u$ and $v$, any vector of the form

$$au + bv$$

is called a *linear combination* of $u$ and $v$. The set of all such combinations is

$$\text{span}(u, v) = \{au + bv : a, b \in \mathbb{R}\}.$$

If $u$ and $v$ are scalar multiples of each other, then $\text{span}(u, v)$ is still just a line. If they point in genuinely different directions, then their span fills out the whole plane.

**Span**

The span of one nonzero vector in $\mathbb{R}^2$ is a line through the origin. The span of two non-parallel vectors in $\mathbb{R}^2$ is the whole plane.

## 1.5 Basis vectors and coordinates

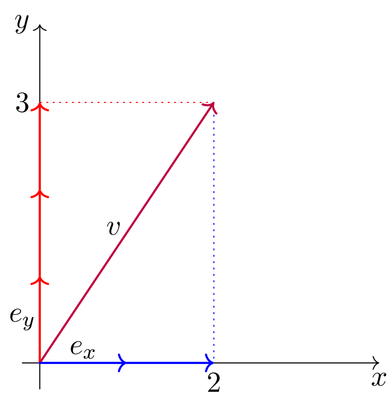
In the plane, two particularly important vectors are

$$e_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

These are called the *standard basis vectors*. In fact, every vector in $\mathbb{R}^2$ can be written as a linear combination of these two vectors. For example:

$$v = \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2e_x + 3e_y.$$

Graphically, we can see this as an arithmetic operation very similar to how we plotted points in 2-dimensional space:




Using these two special vectors, we can reinterpret the components of any vector as:

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$u_1$ ← Amount of movement in the $x$ direction (left/right)

$u_2$ ← Amount of movement in the $y$ direction (up/down)

**Exercise 1**

Let

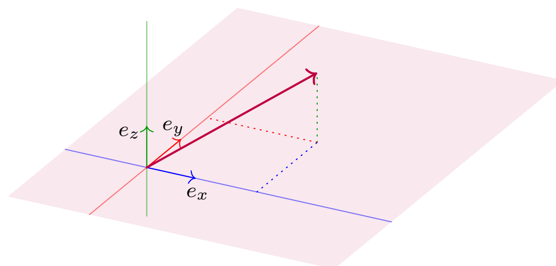
$$u = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \qquad v = \begin{pmatrix} -3 \\ 4 \end{pmatrix}.$$

Compute:

1. $u + v$
2. $2u$
3. $3u - v$

## 1.6 Understanding Higher Dimensions

If we have the two vectors $e_x$ and $e_y$, we can create a two-dimensional plane of vectors. Imagine now that we take this plane, and add another vector of length 1 sticking out of the plane in a different direction:




We will denote this new vector by $e_z$. As you can see from the above, this new vector has a span that is completely perpendicular to both $e_x$ and $e_y$. In other words, $e_z$ gives us access to another *dimension* in which our vectors may be potentially pointing. We may then interpret the 2-dimensional plane spanned by $e_x$ and $e_y$ as the special case in which the $z$ direction is simply zero. In general, we can move this plane around in the third dimension by varying the amount of $e_z$ that we decide to include in any given vector. In order to keep track of this extra bit of information, we can write three-dimensional vectors with *three* numbers instead of the two we used before. Starting with our special basis vectors, we write:

$$e_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad e_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad e_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

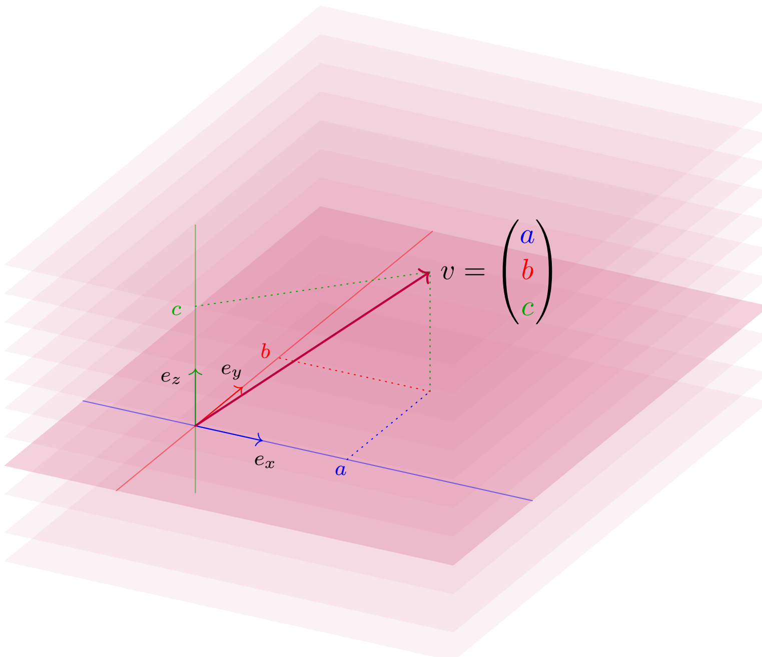
and any vector

$$v = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

may be written as

$$v = ae_x + be_y + ce_z.$$

In a sense, we can imagine the resulting space as a collection of many copies of $\text{span}(e_x, e_y)$ stacked on top of each other:




In the same way that a line consists of an infinite number of points, continuously stacked next to each other, we can understand 3-dimensional space as an infinite collection of 2-dimensional spaces continuously stacked on top of each other.

Although the same concept continues on into higher dimensions, our geometric intuition for the behaviour of vectors quickly breaks down. Instead, we have to rely on algebraic descriptions of vectors. Algebraically, an $n$-dimensional vector has the form

$$v = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n.$$

where the standard basis vectors are

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \cdots \quad e_{n-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

These are $n$ basis vectors pointing in $n$ different perpendicular directions. This vector space of general dimension is written $\mathbb{R}^n$.

**Exercise 2**

1. Write $\begin{pmatrix} 4 \\ -2 \end{pmatrix}$ in the form $ae_x + be_y$.
2. Write $\begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix}$ in the form $ae_x + be_y + ce_z$.
3. Describe $\operatorname{span}\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right)$ geometrically.

# 2 Matrices

We have seen previously that a matrix is a rectangular array of real numbers. Recall that if a matrix has $m$ rows and $n$ columns, then we say that it is an $m \times n$ matrix, or that the *order* of the matrix is $m \times n$.

We will now demonstrate that matrices, in their own way, also admit their own type of arithmetic.

## 2.1 Matrix addition

If we have two matrices with the same order, then we can add them together entry-by-entry. For $2 \times 2$ matrices, addition looks like:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix}.$$

As an example,

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

We will demonstrate this step-by-step:

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1+0 & \\ & \end{pmatrix} = \begin{pmatrix} 1 & \\ & \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1+0 & 2+2 \\ & \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ & \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1+0 & 2+2 \\ 3+2 & \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 5 & \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1+0 & 2+2 \\ 3+2 & 3+1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 5 & 4 \end{pmatrix}$$

It should be emphasized that matrix addition only works when both matrices have the same order, otherwise the entries in each matrix won't "match up" and there will be nothing to add.

## 2.2 Matrix multiplication

Under the right conditions, a pair of matrices can be multiplied together. However, in distinction to matrix addition, this operation is *not* performed entry-by-entry. Instead, matrix multiplication is more subtle. For $2 \times 2$ matrices, we multiply by:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}.$$

Here, we are taking the rows of the first matrix and selectively multiplying them by the columns of the second matrix, for instance, the top-left entry of the multiplication is given by the top row of the first matrix multiplied by the left column of the second matrix. When we multiply a row with a column, we need to make sure that the order of the entries is respected.

As an example, we compute:

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 0 + 2\cdot 2 & 1\cdot 2 + 2\cdot 1 \\ 3\cdot 0 + 3\cdot 2 & 3\cdot 2 + 3\cdot 1 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 6 & 9 \end{pmatrix}.$$

To see how this is performed, here is the step-by-step computation:

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 0 + 2\cdot 2 & \\ & \end{pmatrix} = \begin{pmatrix} 4 & \\ & \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 0 + 2\cdot 2 & 1\cdot 2 + 2\cdot 1 \\ & \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ & \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 0 + 2\cdot 2 & 1\cdot 2 + 2\cdot 1 \\ 3\cdot 0 + 3\cdot 2 & \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 6 & \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 0 + 2\cdot 2 & 1\cdot 2 + 2\cdot 1 \\ 3\cdot 0 + 3\cdot 2 & 3\cdot 2 + 3\cdot 1 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 6 & 9 \end{pmatrix}$$

In contrast to matrix addition, it is also possible to multiply matrices of different orders, the only requirement is that the row length of the left-hand matrix matches the column length of the right-hand matrix.

**Size rule for matrix multiplication**

An $m \times n$ matrix can be multiplied by an $n \times \ell$ matrix, and the result is an $m \times \ell$ matrix.

## 2.3 Zero and identity matrices

Just as numbers have 0 and 1, matrices often also have special neutral elements. For $2 \times 2$ matrices, the zero matrix is

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

and it satisfies $A + 0 = A$ for every $2 \times 2$ matrix. The multiplicative identity matrix of order $2 \times 2$ is:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and it satisfies

$$AI = IA = A$$

for every $2 \times 2$ matrix $A$.

## 2.4 Properties of Matrix Arithmetic

Throughout the course we have seen several instances of arithmetic: we started with a brief discussion of the arithmetic of real numbers, and then much later in the course we began to talk about the arithmetic of polynomials. For the latter, we used the narrative that the polynomials very much behave like the *integers* $\mathbb{Z}$. In this sense, we noted that the collection of all polynomials was not quite like numbers, but they *rhymed*.

Unlike polynomials, matrices do not "rhyme" with ordinary numbers. To begin with, it is not always the case that a pair of matrices can be added or multiplied, their orders must be compatible in specific ways. However, even if this *were* the case, for instance, if we were to restrict our attention to only the square matrices of order $2 \times 2$, then matrix arithmetic still displays some strange behaviours.

**Unusual Properties of Matrix Arithmetic**

For matrices $A$ and $B$:

1. multiplication is not necessarily commutative, so $AB \neq BA$,
2. not every matrix has a multiplicative inverse. In other words, some matrices $A$ have no inverse $A^{-1}$ such that $AA^{-1} = I$.

To see that matrix multiplication does not necessarily commute, consider the pair of $2 \times 2$ matrices:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then, their products $AB$ and $BA$ are not equal:

$$AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \qquad BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

that is, $AB \neq BA$.

**Exercise 3**

Compute each of the following:

1. $\begin{pmatrix} 2 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix}$
2. $\begin{pmatrix} 1 & 2 \\ 1 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 1 & 5 \end{pmatrix}$
3. $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 3 \\ 1 & 1 \end{pmatrix}$
4. $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

**Exercise 4**

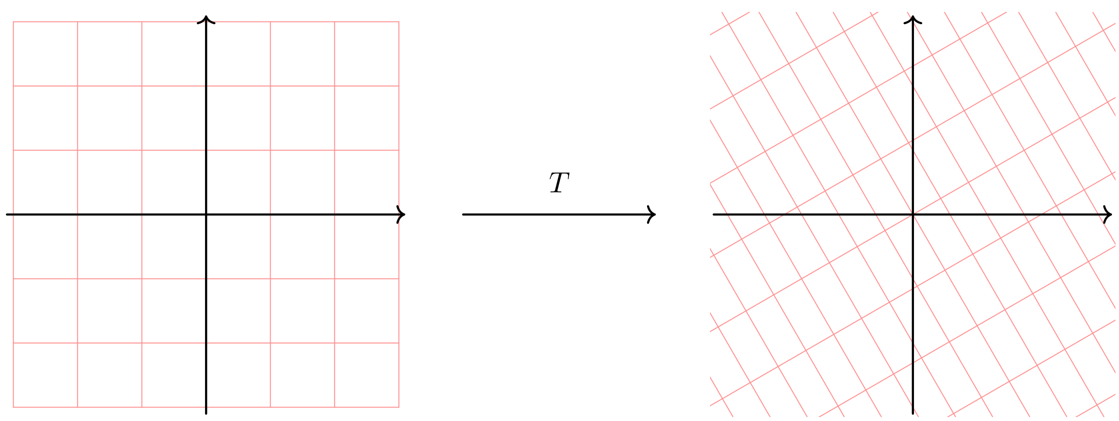
1. Is the product of a $2 \times 3$ matrix and a $3 \times 4$ matrix defined? If so, what is the size of the result?

2. Is the product of a $2 \times 3$ matrix and a $2 \times 2$ matrix defined?
3. Explain in words what the identity matrix does to another matrix when multiplied.

# 3 Linear Transformations

## 3.1 The geometric idea

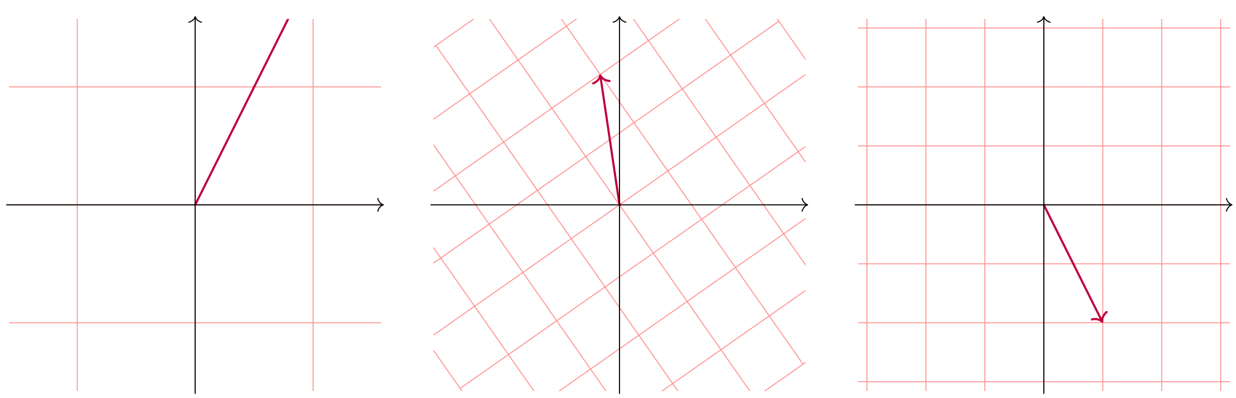
A linear transformation is a function that takes vectors as input and produces vectors as output. Geometrically, linear transformations are the maps that preserve the "straightness" of space: lines stay lines, the square grid in the coordinate plane becomes another grid of straight lines, and the origin stays fixed. For instance, the image below depicts a linear transformation:



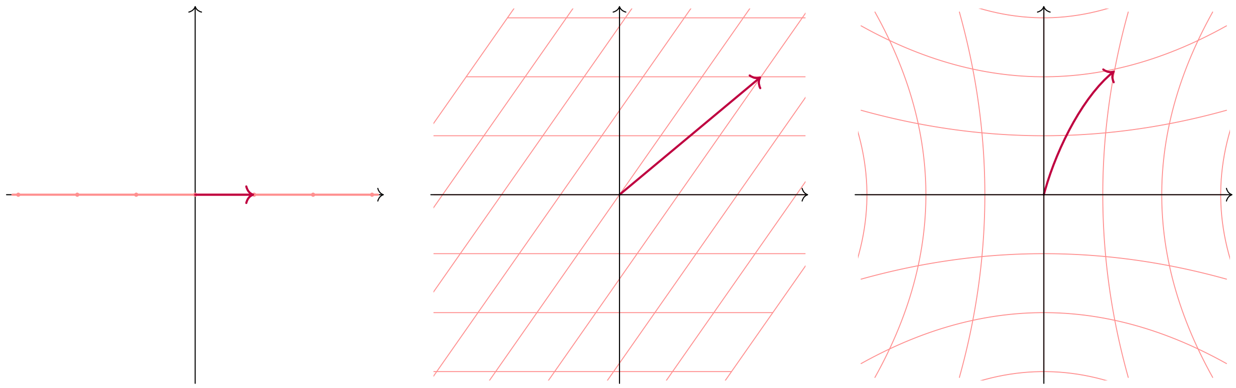

Typical examples include:

- scalings,
- rotations,
- reflections,
- projections, and
- shears.

A map that bends straight lines into curves is not linear.



(i) Scaling by 2 (ii) Rotation (iii) Reflection in the $x$-axis



(iv) Projection onto the $x$-axis

(v) Shearing

(vi) A non-linear transformation

### 3.2 Formal definition

We will focus on transformations of vectors in the plane, so our input and output vectors both lie in $\mathbb{R}^2$. A transformation $T$ is a function $T : \mathbb{R}^2 \to \mathbb{R}^2$, and it is called linear if it preserves vector addition and scalar multiplication:

$$T(u+v) = T(u) + T(v), \qquad T(av) = aT(v).$$

**Definition of a linear transformation**

A function $T$ between vector spaces is linear if, for all vectors $u, v$ and all scalars $a$,

$$T(u+v) = T(u) + T(v) \qquad \text{and} \qquad T(av) = aT(v).$$

These two rules are the algebraic reason that linear transformations preserve the flat grid structure of the plane.

### 3.3 Linear transformations and matrices

A very useful fact is that, in $\mathbb{R}^2$, a linear transformation is completely determined by what it does to the two standard basis vectors

$$e_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Indeed, every vector has the form

$$\begin{pmatrix} x \\ y \end{pmatrix} = xe_x + ye_y.$$

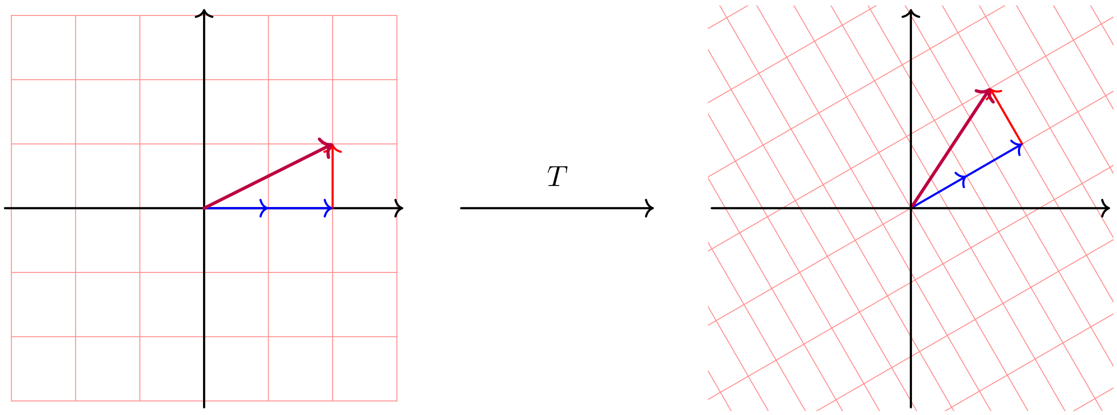
So we can apply the linearity properties of $T$ to conclude that:

$$T\begin{pmatrix} x \\ y \end{pmatrix} = T(xe_x + ye_y) = xT(e_x) + yT(e_y).$$

Geometrically, any linear transformation sends our coordinate grid into another coordinate grid. Since the output grid consists of *straight* lines, we can still navigate and move around the grid using scalar multiplication and addition. The basis vectors $e_x$ and $e_y$ will be transformed into two vectors

$T(e_x)$ and $T(e_y)$ that can be used to "walk" around the new grid and describe where the output of any other vector $T(v)$ will be.

For example, if we have a vector $v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and we would like to know where it lands under some linear transformation $T$, we can look at the output grid and follow two copies of the transformed $e_x$-direction and then one copy of the transformed $e_y$-direction. That endpoint is $T(v)$:




In general, suppose that a linear transformation $T$ sends the basis vectors $e_x$ and $e_y$ to:

$$T(e_x) = \begin{pmatrix} a \\ c \end{pmatrix}, \qquad T(e_y) = \begin{pmatrix} b \\ d \end{pmatrix},$$

where $a, b, c$ and $d$ are real numbers. Since any other vector $v$ can be uniquely written in components:

$$v = \begin{pmatrix} x \\ y \end{pmatrix} = xe_x + ye_y,$$

we can write:

$$T\begin{pmatrix} x \\ y \end{pmatrix} = T\left(xe_x + ye_y\right) = xT(e_x) + yT(e_y) = x\begin{pmatrix} a \\ c \end{pmatrix} + y\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

We now make an important observation: if we interpret the vector $v$ as a $2 \times 1$ *matrix*, then we can perform the matrix multiplication:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix},$$

which is exactly the same as $T(v)$. This means that $T$ is described by the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

whose columns are exactly the outputs of the basis vectors $e_x$ and $e_y$ under the transformation $T$.

Every linear transformation on the plane is a matrix

If

$$T(e_x) = \begin{pmatrix} a \\ c \end{pmatrix}, \qquad T(e_y) = \begin{pmatrix} b \\ d \end{pmatrix},$$

then

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

## 3.4 Examples of transformation matrices

The five examples of linear transformations depicted in Section 3.1 have the following transformation matrices:

(i) scaling by 2:

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

(ii) rotation by angle $\theta$:

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

(iii) reflection in the $x$-axis:

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

(iv) projection onto the $x$-axis:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

(v) shearing:

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

As an explicit computation, if

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}, \qquad v = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

then

$$Av = \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 + 1 \cdot 2 \\ -1 \cdot 1 + 3 \cdot 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

## 3.5 Composition of transformations

Linear transformations map vector spaces to vector spaces. In principle, it is possible to *compose* these together by doing one transformation after another. Given a pair of linear transformations $T$ and $S$, we can write their composition as $S \circ T$. This notation should be read "transform the vector space according to $T$, and then transform the result according to $S$". If $T$ is represented by the matrix $A$ and $S$ is represented by the matrix $B$, then the composition $S \circ T$ is represented by multiplying these two matrices together from the right, that is, the product $BA$. Since matrix multiplication is not commutative, this means that the *order matters*.

To demonstrate this, let's consider an example. Let:

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

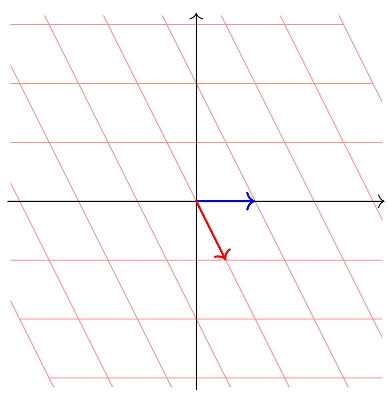
which is a reflection in the $x$-axis, and let

$$B = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix},$$

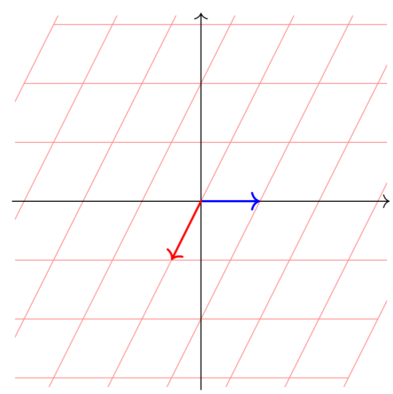
which is a shear. Multiplying these two matrices together in different orders gives:

$$AB = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & -1 \end{pmatrix}, \qquad BA = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & -1 \end{pmatrix},$$

so $AB \neq BA$. Since $A$ is a reflection and $B$ is a shear, the product $BA$ means to first reflect and then shear, while the product $AB$ means to first shear and then reflect. The fact that $AB \neq BA$ means that applying this reflection and then a shear will be different to applying a shear *and then* doing the reflection. Visually, this is very apparent:



Output of the linear transformation $AB$



Output of the linear transformation $BA$

**Exercise 5**

Let

$$T = \begin{pmatrix} 3 & -1 \\ 2 & 4 \end{pmatrix}.$$

Compute:

1. $T\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
2. $T\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
3. $T\begin{pmatrix} 2 \\ -1 \end{pmatrix}$

**Exercise 6**

Suppose a linear transformation $T$ satisfies

$$T(e_x) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad T(e_y) = \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

1. Write down the matrix of $T$.
2. Find $T\left(\begin{pmatrix} 4 \\ 2 \end{pmatrix}\right)$.

# 4 Determinants Revisited

## 4.1 Determinant as area

Recall from the previous lecture that for

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

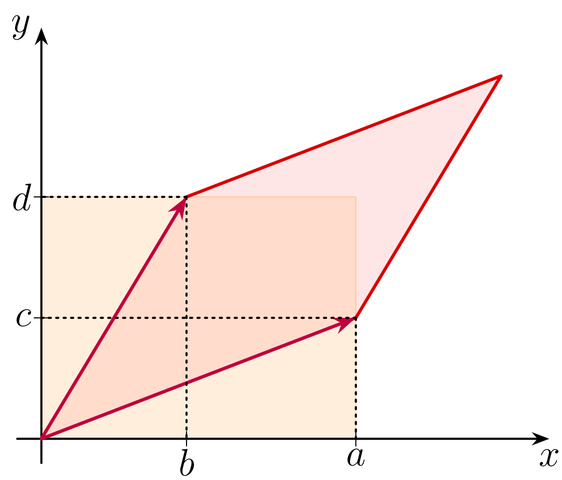
the determinant is

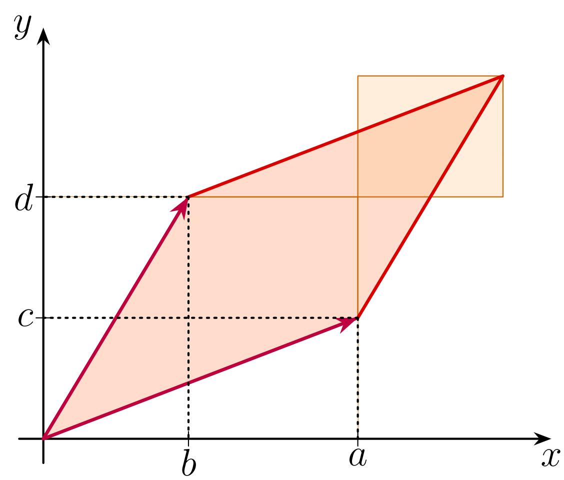
$$\det(A) = ad - bc.$$

If we interpret the columns of $A$ as vectors

$$u = \begin{pmatrix} a \\ c \end{pmatrix}, \qquad v = \begin{pmatrix} b \\ d \end{pmatrix},$$

then $\det(A)$ is closely related to the area of the parallelogram spanned by $u$ and $v$. In the pictured configuration, one can arrange the labels so that a rectangle of area $ad$ contains the parallelogram, while the excess has area $bc$. This can be seen by translating parts of the area $ad$ across the parallelogram:




More precisely,

$$\text{Area of the parallelogram} = \left|\det(A)\right|.$$

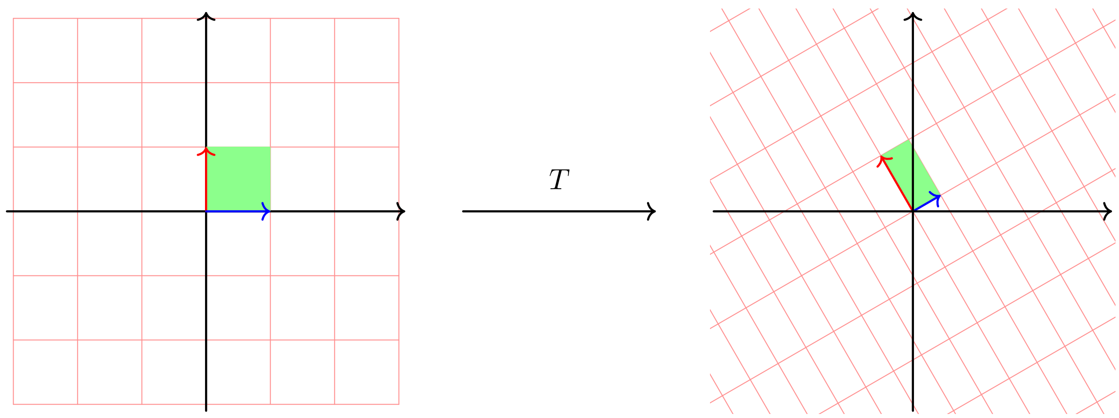
So $2 \times 2$ determinants have a geometric meaning: they measure the area of the parallelograms formed from two vectors.

## 4.2 Determinant as a scale factor

There is another interpretation which is even more important in linear algebra. The matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

describes a linear transformation. Under this transformation, the unit square in the plane is sent to the parallelogram spanned by the columns of $A$. Therefore, $\det(A)$ tells us how area changes under the transformation.




**Determinant as an area scale factor**

If a linear transformation is represented by the matrix $A$, then

$$|\det(A)|$$

tells us by what factor areas are scaled.

For example:

1. If
$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$
then
$$\det(A) = 4.$$
So areas become 4 times larger.
2. If
$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$
then
$$\det(A) = \cos^2\theta + \sin^2\theta = 1.$$
So rotations preserve area.
3. If
$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

then

$$\det(A) = 0.$$

So all areas collapse to zero. This is exactly what we expect from a projection onto a line.

4. If

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

then

$$\det(A) = -1.$$

The absolute value is 1, so area is preserved, but the negative sign indicates that the transformation flips orientation.

## 4.3 A geometric interpretation of $3 \times 3$ determinants

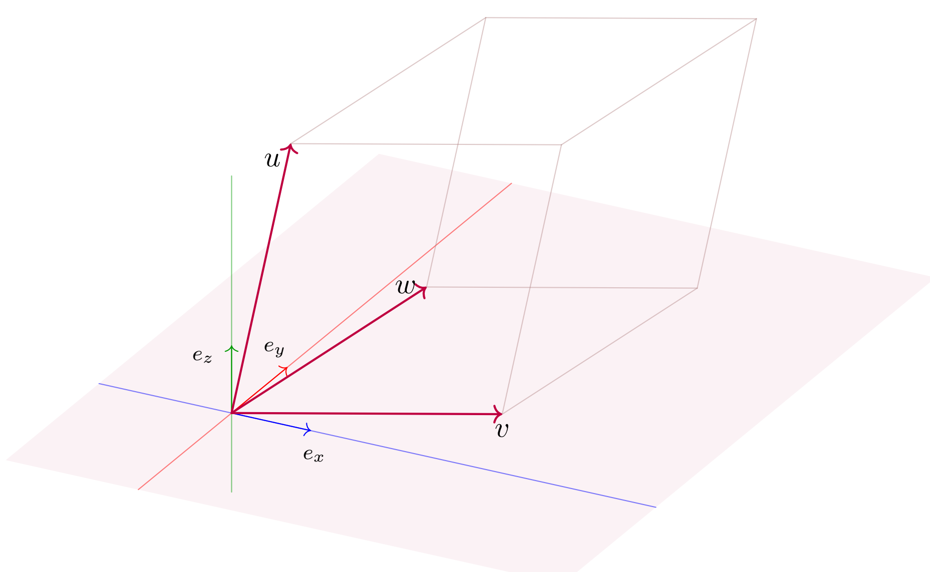
Recall that for a $3 \times 3$ matrix, we may compute the determinant by expansion by minors. For example, along the third row, we would expand by:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

In direct analogy to the 2-dimensional case, the determinant of a $3 \times 3$ matrix provides the signed volume of the parallelepiped formed from the columns, viewed as vectors in 3-dimensional space. Its absolute value gives the ordinary volume, while the sign records orientation. If we let

$$u = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad v = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad \text{and} \quad w = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix},$$

then one example of this parallelepiped may be:

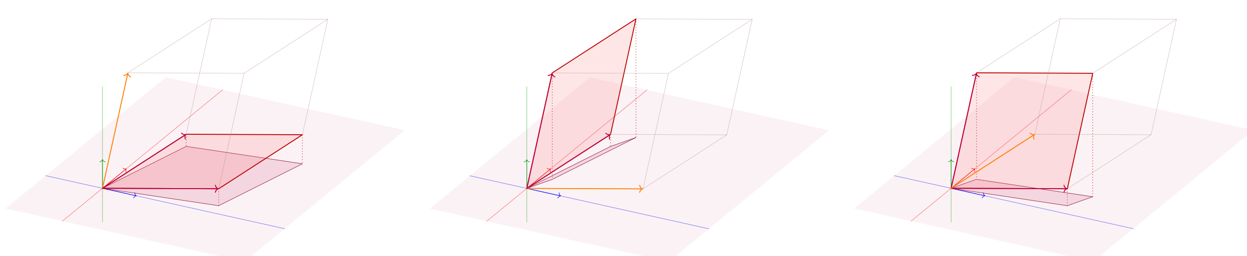


When we expand by the third row, we are breaking down the parallelepiped into a series of $2 \times 2$ areas, which we then multiply by a length to create prisms. In fact, in the case of expansion by the third row, these parallelograms are constructed by taking each of the parallelograms created from a choice of two vectors in $\{u, v, w\}$, forming a parallelogram from these two vectors, and then projecting it down onto the $xy$ plane. The corresponding $2 \times 2$ determinants are signed areas of those projected parallelograms. In annotated form, expansion by the third row computes:

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

- $a_{31}$ is the $z$-coordinate of the vector in column 1
- signed area of the parallelogram formed by columns 2 and 3, after projecting onto the $xy$-plane
- $a_{32}$ is the $z$-coordinate of the vector in column 2
- signed area of the parallelogram formed by columns 1 and 3, after projecting onto the $xy$-plane
- $a_{33}$ is the $z$-coordinate of the vector in column 3
- signed area of the parallelogram formed by columns 1 and 2, after projecting onto the $xy$-plane

These three projected parallelograms are depicted below:



It is a tedious matter of geometry, beyond the scope of this course, to rearrange the associated blocks of parallelogram prisms into the exact volume of the parallelepiped. However, this is indeed possible.

**Exercise 7**

1. Find the determinant of $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$. What does this tell us about area scaling?
2. Find the determinant of $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. What happens to area?
3. Compute $\begin{vmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \\ 2 & 0 & 1 \end{vmatrix}$ by expanding along the first row.

## 5 Matrices and Solution Spaces

Linear algebra is deeply connected to systems of equations. A system such as

$$\begin{cases} 3x + 2y - z = 1, \\ x + 2z = -3, \\ -2x - y = 4 \end{cases}$$

can be rewritten as a matrix multiplication:

$$\begin{pmatrix} 3 & 2 & -1 \\ 1 & 0 & 2 \\ -2 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix}.$$

In this sense, we can understand the system as asking for the input vector that, once transformed according to the coefficient matrix, yields the output vector.
This compact notation is one of the main reasons matrices are so useful. Instead of writing several equations one after another, we can present the whole system in a single matrix equation

$$A\mathbf{x} = \mathbf{b}.$$

For a general system, it is better to speak of the *solution set*: the set of all vectors $\mathbf{x}$ that satisfy this equation. As we already saw in earlier lectures, there are only three possibilities:

1. one solution,
2. infinitely many solutions,
3. no solution.

For square systems, determinants often help us detect the first case quickly. Analogous to Cramer's rule that we covered in Lecture 16, here we have the following.

**Determinant and uniqueness**

If $A$ is a square coefficient matrix and $\det(A) \neq 0$, then the system

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

has a unique solution.
If $\det(A) = 0$, then the system does not have a unique solution. It may have infinitely many solutions or no solution.

### Exercise 8

Use determinants or simple inspection to decide whether each system has a unique solution.

1. $\begin{cases} x + y = 2 \\ x - y = 0 \end{cases}$
2. $\begin{cases} x + y = 1 \\ 2x + 2y = 2 \end{cases}$

3. $\begin{cases} x + y = 1 \\ 2x + 2y = 3 \end{cases}$

---

## Solutions to the Exercises

### Exercise 1

$$u + v = \begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} -3 \\ 4 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

$$2u = 2 \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}.$$

$$3u - v = 3 \begin{pmatrix} 2 \\ -1 \end{pmatrix} - \begin{pmatrix} -3 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \end{pmatrix} - \begin{pmatrix} -3 \\ 4 \end{pmatrix} = \begin{pmatrix} 9 \\ -7 \end{pmatrix}.$$

### Exercise 2

1. $$\begin{pmatrix} 4 \\ -2 \end{pmatrix} = 4e_x - 2e_y.$$

2. $$\begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} = 1e_x + 0e_y + 5e_z = e_x + 5e_z.$$

3. $$\text{span}\left( \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right)$$

   is the line through the origin in the direction $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

### Exercise 3

1. $$\begin{pmatrix} 2 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 6 & 6 \end{pmatrix}.$$

2. $$\begin{pmatrix} 1 & 2 \\ 1 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 1 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 2 & 10 \end{pmatrix}.$$

3. $$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 3 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 + 1 \cdot 1 & 1 \cdot 3 + 1 \cdot 1 \\ 1 \cdot 0 + 2 \cdot 1 & 1 \cdot 3 + 2 \cdot 1 \end{pmatrix}$$
   $$= \begin{pmatrix} 1 & 4 \\ 2 & 5 \end{pmatrix}.$$

4.

$$\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}.$$

### Exercise 4

1. Yes. A $2 \times 3$ matrix times a $3 \times 4$ matrix is defined, and the result is a $2 \times 4$ matrix.
2. No. The inside dimensions do not match, since $3 \neq 2$.
3. Multiplying by the identity matrix leaves the matrix unchanged.

### Exercise 5

1.

$$T\begin{pmatrix}1\\ 0\end{pmatrix} = \begin{pmatrix}3\\ 2\end{pmatrix}.$$

2.

$$T\begin{pmatrix}0\\ 1\end{pmatrix} = \begin{pmatrix}-1\\ 4\end{pmatrix}.$$

3.

$$\begin{aligned} T\begin{pmatrix}2\\ -1\end{pmatrix} &= \begin{pmatrix}3 & -1\\ 2 & 4\end{pmatrix}\begin{pmatrix}2\\ -1\end{pmatrix} \\ &= \begin{pmatrix}3 \cdot 2 + (-1)(-1)\\ 2 \cdot 2 + 4(-1)\end{pmatrix} = \begin{pmatrix}7\\ 0\end{pmatrix}. \end{aligned}$$

### Exercise 6

1. Since the columns are $T(e_x)$ and $T(e_y)$, the matrix is

$$\begin{pmatrix}2 & -1\\ 1 & 3\end{pmatrix}.$$

2.

$$\begin{aligned} T\begin{pmatrix}4\\ 2\end{pmatrix} &= \begin{pmatrix}2 & -1\\ 1 & 3\end{pmatrix}\begin{pmatrix}4\\ 2\end{pmatrix} \\ &= \begin{pmatrix}8 - 2\\ 4 + 6\end{pmatrix} = \begin{pmatrix}6\\ 10\end{pmatrix}. \end{aligned}$$

### Exercise 7

1.

$$\det\begin{pmatrix}2 & 0\\ 0 & 3\end{pmatrix} = 2 \cdot 3 - 0 \cdot 0 = 6.$$

So areas scale by a factor of 6.

2.

$$\det\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = 1(-1) - 0 = -1.$$

The absolute value is 1, so area is preserved, although the plane is flipped.

3.

$$\begin{vmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \\ 2 & 0 & 1 \end{vmatrix} = 1\begin{vmatrix} 3 & 1 \\ 0 & 1 \end{vmatrix} - 2\begin{vmatrix} 0 & 1 \\ 2 & 1 \end{vmatrix} + 0\begin{vmatrix} 0 & 3 \\ 2 & 0 \end{vmatrix}$$

$$= 1(3) - 2(0-2) + 0$$

$$= 3 + 4 = 7.$$

## Exercise 8

1. The coefficient matrix is

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad \det = -1 - 1 = -2 \neq 0.$$

   So the system has a unique solution.

2. The coefficient matrix is

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, \qquad \det = 2 - 2 = 0.$$

   The second equation is just twice the first, so the system has infinitely many solutions.

3. The coefficient matrix is again

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, \qquad \det = 0.$$

   But now the equations are inconsistent, because doubling the first equation gives $2x + 2y = 2$, not $2x + 2y = 3$. So the system has no solution.