Bachelor’s degree in Bioinformatics

Vector and Matrix models

Prof. Renato Bruni

bruni@diag.uniroma1.it

Department of Computer, Control, and Management Engineering (DIAG)

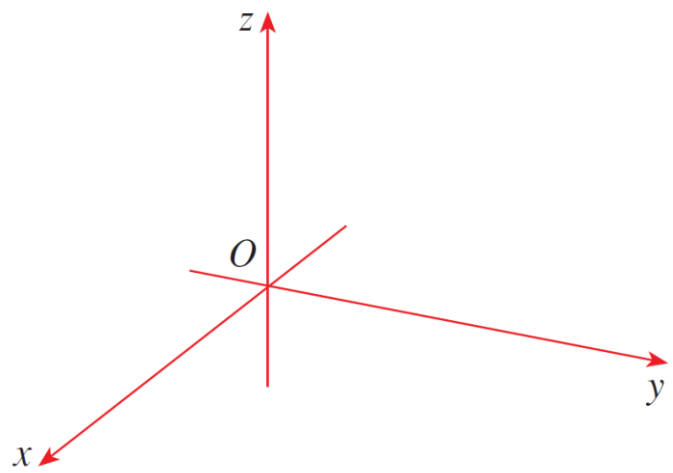
“Sapienza” University of Rome

Three-Dimensional Space

To locate a point in space, three numbers are required. We can represent any point in space by an ordered triple (a, b, c) of real numbers.

In order to do so we first choose a fixed point O (the origin) and three directed lines through O that are perpendicular to each other, called the **coordinate axes** and labeled the x-axis, y-axis, and z-axis.

Usually we think of the x- and y-axes as being horizontal and the z-axis as being vertical, and we draw the orientation of the axes

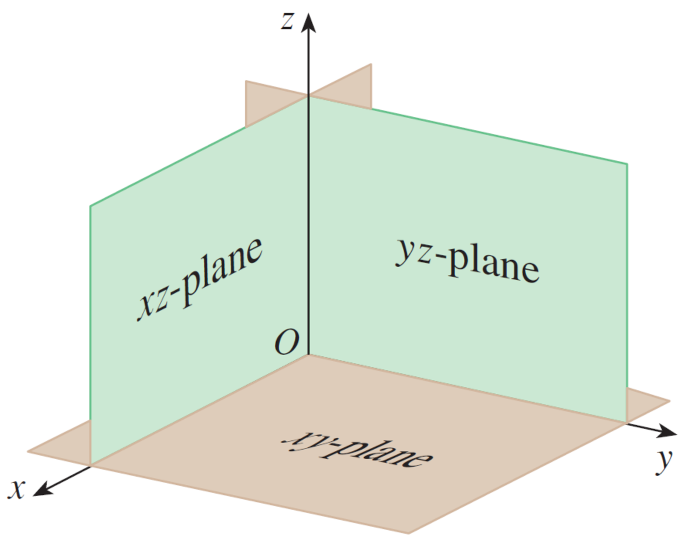


Coordinate axes

Three-Dimensional Space

The three coordinate axes determine also the three **coordinate planes**

The *xy*-plane is the plane that contains the *x*- and *y*-axes; the *yz*-plane contains the *y*- and *z*-axes; the *xz*-plane contains the *x*- and *z*-axes.

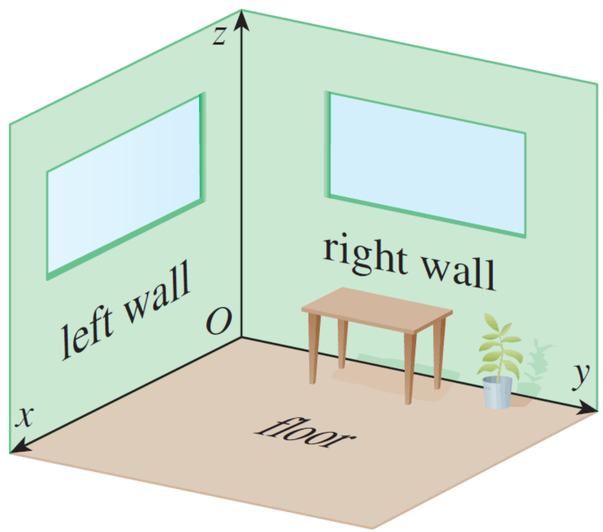


(a) Coordinate planes

Three-Dimensional Space

These three coordinate planes divide space into eight parts, called **octants**. The **first octant**, in the foreground, is determined by the positive axes.

Example: a room with coordinate axes



Now take a point *P* in the space, let:

a be the (directed) distance from the *yz*-plane to *P*,

b the distance from *xz*-plane to *P*,

c the distance from the *xy*-plane to *P*

Three-Dimensional Space

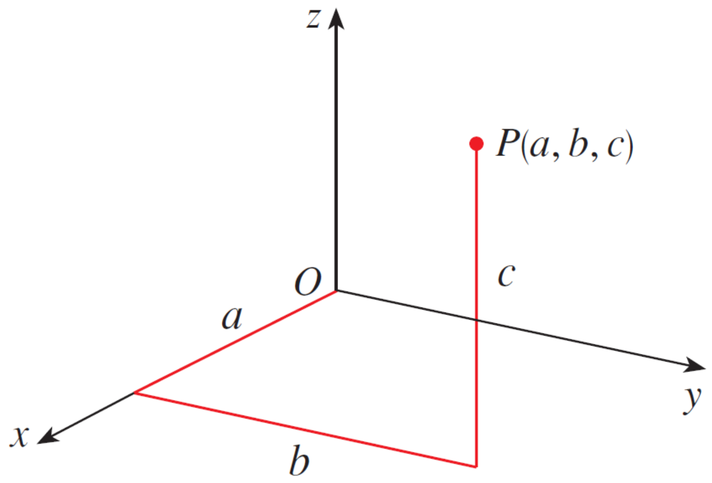
We represent the point *P* by the **ordered triple** (*a*, *b*, *c*) of real numbers and we call *a*, *b*, *c* the **coordinates** of *P*

a is the *x*-coordinate,

b is the *y*-coordinate,

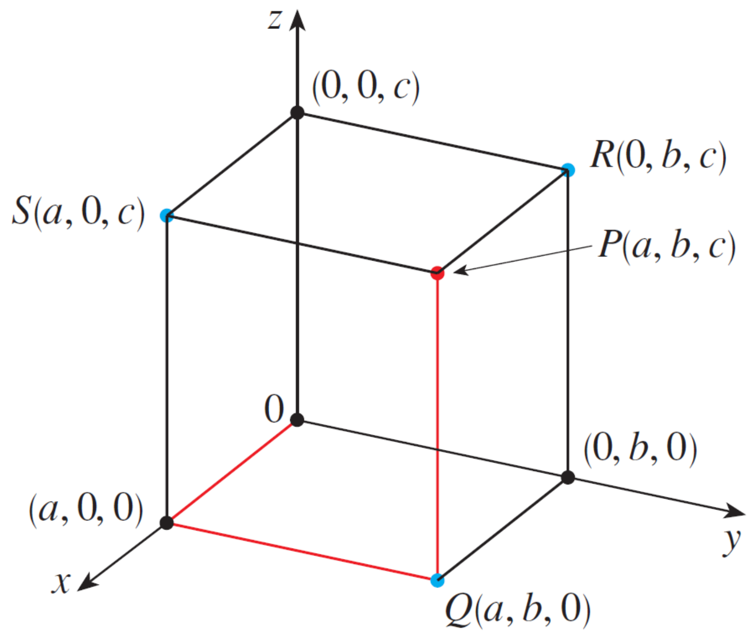
c is the *z*-coordinate.

Thus, to locate the point (*a*, *b*, *c*), we can start at the origin *O* and move *a* units along the *x*-axis, then *b* units parallel to the *y*-axis, and then *c* units parallel to the *z*-axis



Three-Dimensional Space

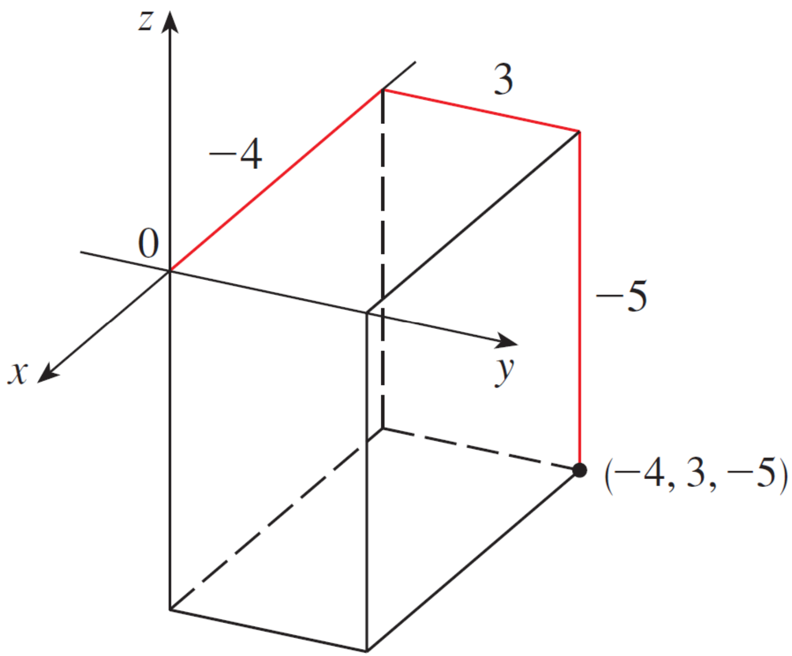
The point $P(a, b, c)$ determines a rectangular box
If we drop a perpendicular from P to the xy-plane, we get a point Q with coordinates $(a, b, 0)$ called the **projection** of P onto the xy-plane.

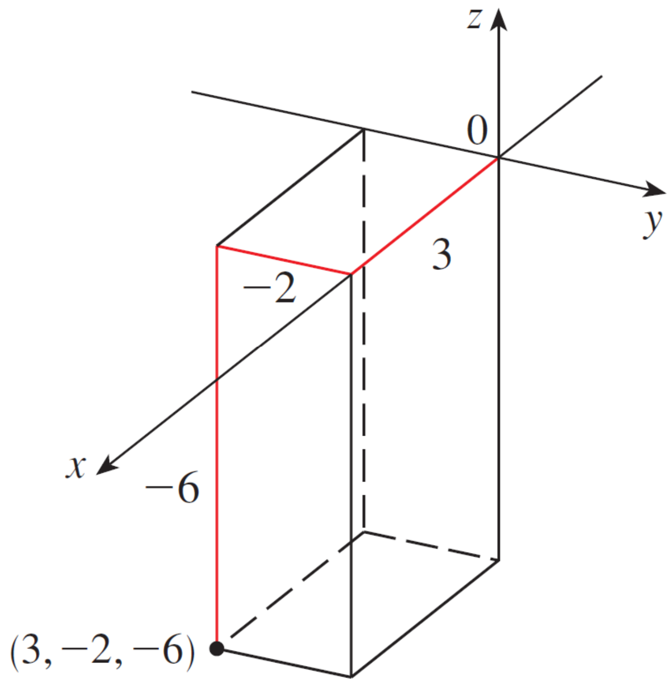


Similarly, $R(0, b, c)$ and $S(a, 0, c)$ are the projections of P onto the yz-plane and xz-plane, respectively.

Three-Dimensional Space

We plot for example the points (–4, 3, –5) and (3, –2, –6)





Three-Dimensional Space

The Cartesian product $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$ is the set of all ordered triples of real numbers and is denoted by $\mathbb{R}^3$.

We have a one-to-one correspondence between points *P* in space and ordered triples (*a*, *b*, *c*) in $\mathbb{R}^3$.

This set is called **space of real numbers of dimension 3**. Such a system of identifying points is called a **three-dimensional rectangular coordinate system**. Notice that, in terms of coordinates, the first octant can be described as the set of points whose coordinates are all positive.

Distance between 2 points

The familiar formula for the distance between two points in a plane is easily extended to the following three-dimensional formula.

Distance Formula in Three Dimensions The distance $|P_1P_2|$ between the points $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ is

$$|P_1P_2| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

Example

The distance from the point $P(2, -1, 7)$ to the point $Q(1, -3, 5)$ is

$$|PQ| = \sqrt{(1 - 2)^2 + (-3 + 1)^2 + (5 - 7)^2}$$

$$= \sqrt{1 + 4 + 4}$$

$$= 3$$

Example of distance

Equation of a Sphere An equation of a sphere with center $C(h, k, l)$ and radius r is

$$(x - h)^2 + (y - k)^2 + (z - l)^2 = r^2$$

In particular, if the center is the origin O, then an equation of the sphere is

$$x^2 + y^2 + z^2 = r^2$$

Higher-Dimensional Space

Just as we extended two-dimensional space to three-dimensional space, we can go further and generalize to **n-dimensional** space.

Although we cannot visualize such spaces, we can still work with them mathematically.

This is extremely important in the life sciences because the systems and objects that we seek to describe mathematically are often characterized by several dimensions.

Higher-Dimensional Space

The Cartesian product $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$, where the product involves *n* copies of $\mathbb{R}$, is denoted $\mathbb{R}^n$ and is defined as $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \in \mathbb{R}\}$.

This is the set of all ordered *n*-tuples of real numbers.

Distance Formula in *n* Dimensions The distance $|P_1P_2|$ between the points $P_1(a_1, \ldots, a_n)$ and $P_2(b_1, \ldots, b_n)$ is

$$|P_1P_2| = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}$$

Vectors

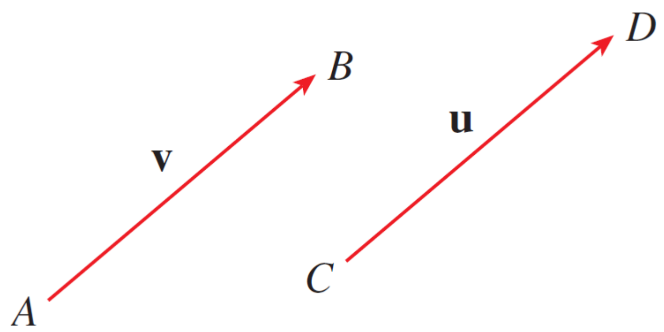
The term **vector** is used by scientists to indicate a quantity (such as displacement or velocity or force) that has both magnitude and direction.

A vector is often represented by an arrow or a directed line segment. The length of the arrow represents the magnitude of the vector and the arrow points in the direction of the vector.

We sometimes denote a vector with a lowercase boldface letter (**v**) or with an arrow above a letter $(\vec{v})$.

Vectors

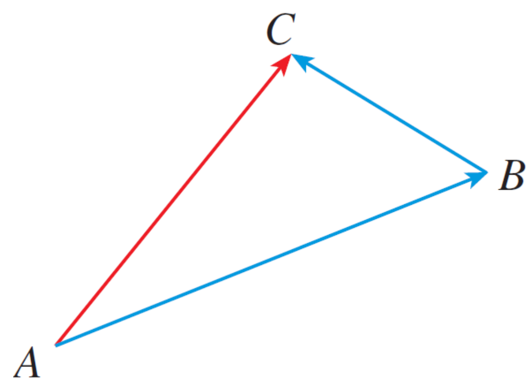
For instance, suppose a particle moves along a line from point *A* to point *B*. The corresponding **displacement vector v** has **initial point** *A* (the tail) and **terminal point** *B* (the tip) and we write $\mathbf{v} = \overrightarrow{AB}.$



Notice that the vector **u** has the same length and the same direction as **v** even though it is in a different position.
We say that **u** and **v** are **equivalent** and we write **u** = **v**.
The **zero vector**, denoted by **0**, has length 0. It is the only vector with no specific direction.

Combining Vectors

Suppose a particle moves from *A* to *B*, so its displacement vector is $\overrightarrow{AB}$. Then the particle changes direction and moves from *B* to *C*, with displacement vector $\overrightarrow{BC}$



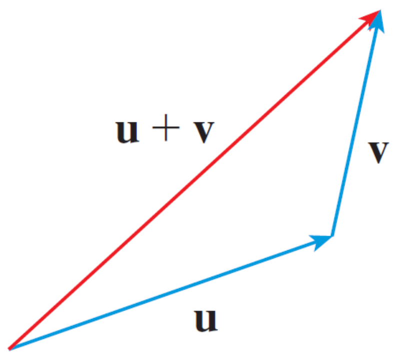
The combined effect of these displacements is that the particle has moved from *A* to *C*. The resulting displacement vector $\overrightarrow{AC}$ is called the *sum* of $\overrightarrow{AB}$ and $\overrightarrow{BC}$ and we write

$$\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$$

Combining Vectors

Definition of Vector Addition If **u** and **v** are vectors positioned so the initial point of **v** is at the terminal point of **u**, then the sum **u** + **v** is the vector from the initial point of **u** to the terminal point of **v**.

The definition of vector addition is illustrated here.

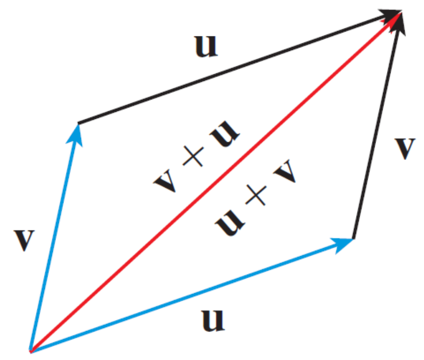


The Triangle Law

This definition is sometimes called the **Triangle Law**.

Combining Vectors

we can start with the same vectors **u** and **v** and draw another copy of **v** with the same initial point as **u**.



The Parallelogram Law

Completing the parallelogram, we see that **u** + **v** = **v** + **u**. This also gives another way to construct the sum: If we place **u** and **v** so they start at the same point, then **u** + **v** lies along the diagonal of the parallelogram with **u** and **v** as sides. This is called the **Parallelogram Law**.

Combining Vectors

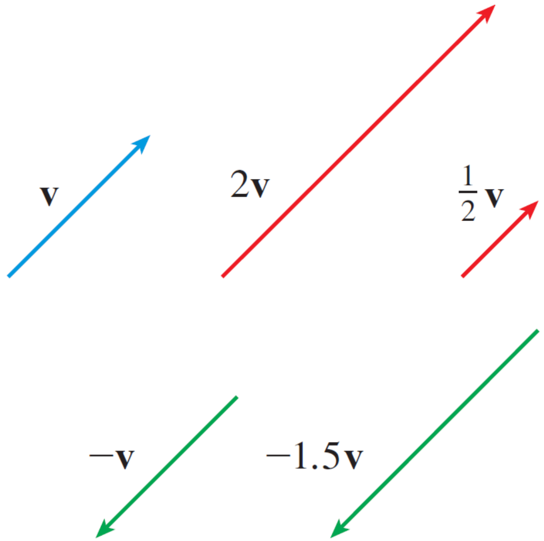
It is possible to multiply a vector by a real number *c*. (In this context we call the real number *c* a **scalar** to distinguish it from a vector.)

For instance, we want 2**v** to be the same vector as **v** + **v**, which has the same direction as **v** but is twice as long. In general, we multiply a vector by a scalar as follows.

Definition of Scalar Multiplication If c is a scalar and $\mathbf{v}$ is a vector, then the **scalar multiple** $c\mathbf{v}$ is the vector whose length is $|c|$ times the length of $\mathbf{v}$ and whose direction is the same as $\mathbf{v}$ if $c > 0$ and is opposite to $\mathbf{v}$ if $c < 0$. If $c = 0$ or $\mathbf{v} = \mathbf{0}$, then $c\mathbf{v} = \mathbf{0}$.

Combining Vectors

This definition is illustrated here



Scalar multiples of **v**

Real numbers work like scaling factors here; that's why we call them scalars. Notice that two nonzero vectors are **parallel** if they are scalar multiples of one another.

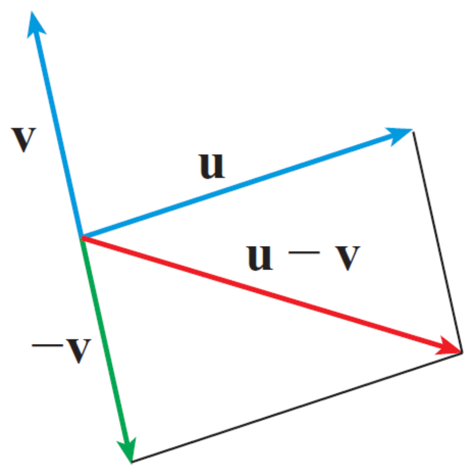
Combining Vectors

In particular, the vector –**v** = (–1)**v** has the same length as **v** but points in the opposite direction. We call it the **negative** of **v**.

By the **difference u – v** of two vectors we mean

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$$

So we can construct **u – v** by first drawing the negative of **v**, –**v**, and then adding it to **u** by the Parallelogram Law

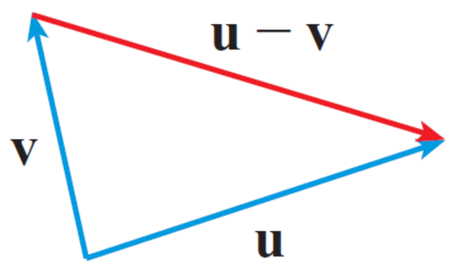


Drawing **u – v**

Combining Vectors

Alternatively, since **v** + (**u** – **v**) = **u**, the vector **u** – **v**, when added to **v**, gives **u**.

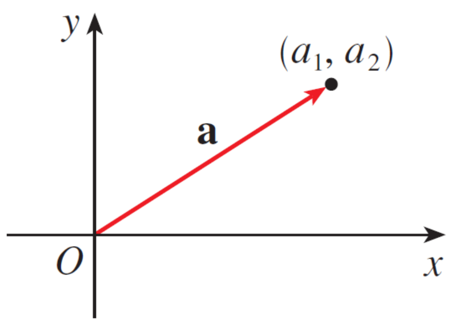
So we could construct **u** – **v** by means of the Triangle Law.



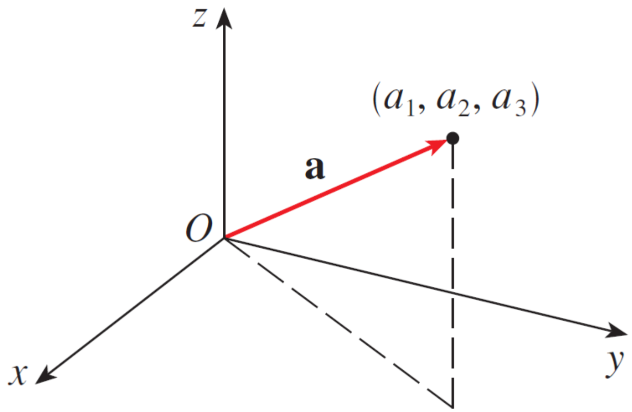
Drawing **u** – **v**

Components

From the mathematical point of view, it is more convenient to introduce a coordinate system and treat vectors algebraically. If we place the initial point of a vector **a** at the origin of a rectangular coordinate system, then the terminal point of **a** has coordinates of the form (a_1, a_2) or (a_1, a_2, a_3), depending on whether we are in 2 or 3 dimensions



$\mathbf{a} = [a_1, a_2]$



$\mathbf{a} = [a_1, a_2, a_3]$

Components

These coordinates are called the **components** of **a** and we write

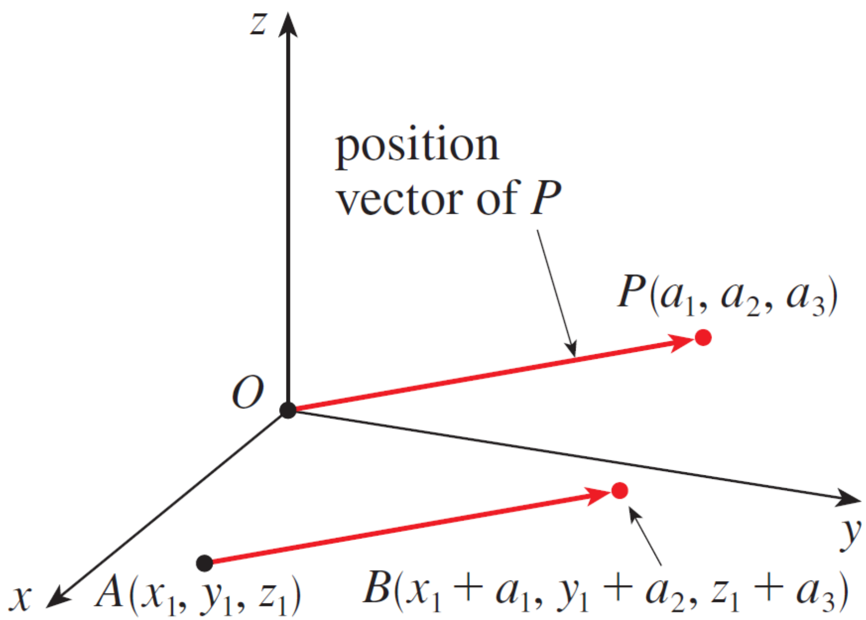
$$\mathbf{a} = [a_1, a_2] \qquad \text{or} \qquad \mathbf{a} = [a_1, a_2, a_3]$$

This representation of vectors with **components** is much more convenient than drawing the vectors as arrows to do every calculation.

We may use the notation $[a_1, a_2]$ for the ordered pair that refers to a vector, so as not to confuse it with the ordered pair (a_1, a_2) that refers to a point in the plane. However, there is no structural differences.

Position Vector

In three dimensions, the vector **a** = $\overrightarrow{OP}$ = $[a_1, a_2, a_3]$ is the **position vector** of the point $P(a_1, a_2, a_3)$



Representations of **a** = $[a_1, a_2, a_3]$

Components

Let's consider any other representation $\overrightarrow{AB}$ of **a**, where the initial point is $A(x_1, y_1, z_1)$ and the terminal point is $B(x_2, y_2, z_2)$. Then we must have

$x_2 = x_1 + a_1$, $y_2 = y_1 + a_2$, and $z_2 = z_1 + a_3$ and so

$a_1 = x_2 - x_1$,

$a_2 = y_2 - y_1$,

$a_3 = z_2 - z_1$

Thus we have the following result.

(1) Given the points $A(x_1, y_1, z_1)$ and $B(x_2, y_2, z_2)$, the vector **a** with representation $\overrightarrow{AB}$ is

$$\mathbf{a} = [x_2 - x_1, y_2 - y_1, z_2 - z_1]$$

Magnitude

The **module**, or **magnitude,** or **length** of a vector **a** is denoted by the symbol $|\mathbf{a}|$ or $\|\mathbf{a}\|$. It is computed by using the distance formula

The length of the two-dimensional vector $\mathbf{a} = [a_1, a_2]$ is

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2}$$

The length of the three-dimensional vector $\mathbf{a} = [a_1, a_2, a_3]$ is

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

A **unit vector** is a vector whose length is 1. In general, if $\mathbf{a} \neq \mathbf{0}$, then the unit vector that has the same direction as **a** is

$$\mathbf{u} = \frac{1}{|\mathbf{a}|}\mathbf{a} = \frac{\mathbf{a}}{|\mathbf{a}|}$$

Operations using Components

If $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$, then

$$\mathbf{a} + \mathbf{b} = [a_1 + b_1, a_2 + b_2] \qquad \mathbf{a} - \mathbf{b} = [a_1 - b_1, a_2 - b_2]$$

$$c\mathbf{a} = [ca_1, ca_2]$$

Similarly, for three-dimensional vectors,

$$[a_1, a_2, a_3] + [b_1, b_2, b_3] = [a_1 + b_1, a_2 + b_2, a_3 + b_3]$$

$$[a_1, a_2, a_3] - [b_1, b_2, b_3] = [a_1 - b_1, a_2 - b_2, a_3 - b_3]$$

$$c[a_1, a_2, a_3] = [ca_1, ca_2, ca_3]$$

Example

If **a** = [4, 0, 3] and **b** = [–2, 1, 5], find |**a**| and the vectors **a** + **b**, **a** – **b**, 3**b**, and 2**a** + 5**b**.

Solution:

$$|\mathbf{a}| = \sqrt{4^2 + 0^2 + 3^2} = \sqrt{25} = 5$$

$$\mathbf{a} + \mathbf{b} = [4, 0, 3] + [-2, 1, 5]$$

$$= [4 + (-2), 0 + 1, 3 + 5] \quad = [2, 1, 8]$$

$$\mathbf{a} - \mathbf{b} = [4, 0, 3] - [-2, 1, 5]$$

$$= [4 - (-2), 0 - 1, 3 - 5] \quad = [6, -1, -2]$$

Example – *Solution*

cont'd

$$3\mathbf{b} = 3[-2, 1, 5] = [3(-2), 3(1), 3(5)]$$

$$= [-6, 3, 15]$$

$$2\mathbf{a} + 5\mathbf{b} = 2[4, 0, 3] + 5[-2, 1, 5]$$

$$= [8, 0, 6] + [-10, 5, 25]$$

$$= [-2, 5, 31]$$

Components

We denote by V_2 the set of all two-dimensional vectors and by V_3 the set of all three-dimensional vectors. More generally, V_n is the set of all n-dimensional vectors where an n-dimensional vector is an ordered n-tuple:

$$\mathbf{a} = [a_1, a_2, \ldots, a_n]$$

and $a_1, a_2, \ldots, a_n$ are real numbers called the components of **a**. Addition and scalar multiplication are defined in terms of components just as for the cases $n = 2$ and $n = 3$.

Properties of vectors

Likewise, the length of a vector from V_n is calculated by using the distance formula.

Properties of Vectors If $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are vectors in V_n and c and d are scalars, then

1. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
2. $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$
3. $\mathbf{a} + \mathbf{0} = \mathbf{a}$
4. $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$
5. $c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}$
6. $(c + d)\mathbf{a} = c\mathbf{a} + d\mathbf{a}$
7. $(cd)\mathbf{a} = c(d\mathbf{a})$
8. $1\mathbf{a} = \mathbf{a}$

The Dot Product

(1) Definition If $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$, then the **dot product** of $\mathbf{a}$ and $\mathbf{b}$ is the number $\mathbf{a} \cdot \mathbf{b}$ given by

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

Thus, to find the dot product of **a** and **b**, we multiply corresponding components and add.

The result is not a vector. It is a real number, that is, a scalar. For this reason, the dot product is sometimes called the **scalar product**.

The Dot Product

The dot product can be defined in a similar fashion for two-dimensional vectors:

$$[a_1, a_2] \cdot [b_1, b_2] = a_1 b_1 + a_2 b_2$$

Likewise, for *n*-dimensional vectors we have:

$$[a_1, \ldots, a_n] \cdot [b_1, \ldots, b_n] = a_1 b_1 + \cdots + a_n b_n$$

Example

$$[2, 4] \cdot [3, -1] = 2(3) + 4(-1)$$
$$= 2$$

$$[-1, 7, 4] \cdot \left[6, 2, -\tfrac{1}{2}\right] = (-1)(6) + 7(2) + 4\left(-\tfrac{1}{2}\right)$$
$$= 6$$

$$[1, 2, -3] \cdot [0, 2, -1] = 1(0) + 2(2) + (-3)(-1)$$
$$= 7$$

The Dot Product

The dot product obeys many of the laws that hold for ordinary products of real numbers.

(2) Properties of the Dot Product If **a**, **b**, and **c** are vectors in V_3 and c is a scalar, then

1. $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$
2. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$
3. $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$
4. $(c\mathbf{a}) \cdot \mathbf{b} = c(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (c\mathbf{b})$
5. $\mathbf{0} \cdot \mathbf{a} = 0$

(4) An Alternative Formula for the Dot Product

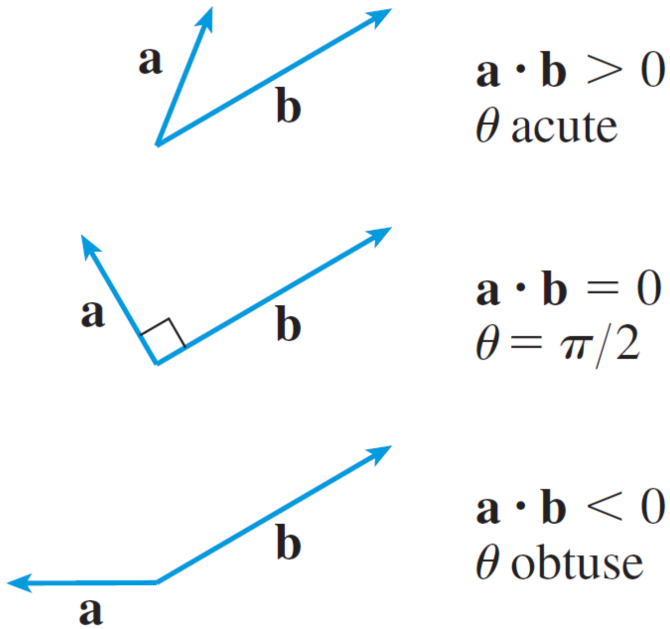
$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$$

where θ is the angle between **a** and **b** ($0 \leqslant \theta \leqslant \pi$). ($\theta$ is the smaller angle between the two vectors when drawn from the same initial point.)

Orthogonal vectors

Notice that, if the dot product of two nonzero vectors is zero, then must be $\cos\theta = 0$

Thus $\theta = \pi/2$ and the two vectors are perpendicular, or **orthogonal**.



Sign of the Dot Product

Also, because cos $\theta > 0$ if $0 \leq \theta < \pi/2$ and cos $\theta < 0$ if $\pi/2 < \theta \leq \pi$, we see that $\mathbf{a} \cdot \mathbf{b}$ is positive for $\theta < \pi/2$ and negative for $\theta > \pi/2$.

We can therefore think of $\mathbf{a} \cdot \mathbf{b}$ as measuring the extent to which **a** and **b** point in the same direction.

The dot product $\mathbf{a} \cdot \mathbf{b}$ is positive if **a** and **b** point in the same general direction, 0 if they are perpendicular, and negative if they point in generally opposite directions.

Same or opposite directions

In the extreme case where **a** and **b** point in exactly the same direction we have $\theta = 0$, so cos $\theta = 1$ and

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|$$

If **a** and **b** point in exactly opposite directions, then $\theta = \pi$ and so cos $\theta = -1$ and $\mathbf{a} \cdot \mathbf{b} = -|\mathbf{a}||\mathbf{b}|$.

Example

Find the angle between the vectors **a** = [2, 2, –1] and **b** = [5, –3, 2].

Solution:

Let θ be the required angle. Since

$$|\mathbf{a}| = \sqrt{2^2 + 2^2 + (-1)^2} = 3 \quad \text{and} \quad |\mathbf{b}| = \sqrt{5^2 + (-3)^2 + 2^2} = \sqrt{38}$$

and since

$$\mathbf{a} \cdot \mathbf{b} = 2(5) + 2(-3) + (-1)(2) = 2$$

we obtain

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{2}{3\sqrt{38}}$$

Example – *Solution*

cont'd

So the angle between **a** and **b** is

$$\theta = \cos^{-1}\left(\frac{2}{3\sqrt{38}}\right) \approx 1.46 \quad (\text{or } 84^\circ)$$

Plane orthogonal to a vector

A plane is determined by a point $P_0(x_0, y_0, z_0)$ in the plane and a vector **n** = $[a, b, c]$ that is orthogonal to the plane. This orthogonal vector **n** is called a **normal vector**.

(5) An equation of the plane that passes through the point $P_0(x_0, y_0, z_0)$ and is perpendicular to the vector $[a, b, c]$ is

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

Example

Find an equation of the plane through the point (2, 4, –1) with normal vector **n** = [2, 3, 4].

Solution:

Putting $a = 2$, $b = 3$, $c = 4$, $x_0 = 2$, $y_0 = 4$, and $z_0 = -1$ in Equation 5, we see that an equation of the plane is

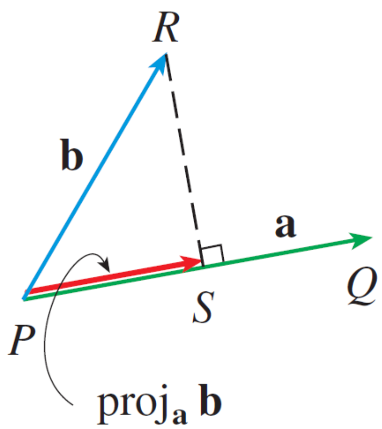
$$2(x - 2) + 3(y - 4) + 4(z + 1) = 0$$

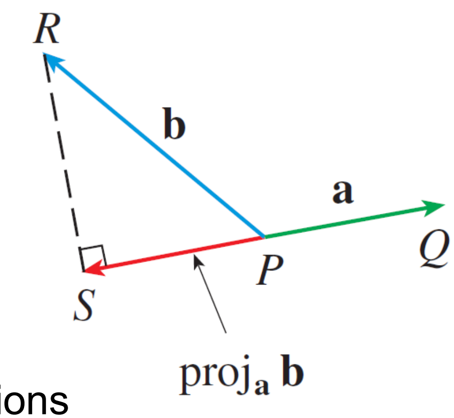
or

$$2x + 3y + 4z = 12$$

Projections

Figure shows representations of two arbitrary vectors **a** and **b** with the same initial point *P*. If *S* is the base of the perpendicular from *R* to the line containing **a**, then the vector with representation $\overrightarrow{PS}$ is called the **vector projection** of **b** onto **a** and is denoted by proj_a **b**

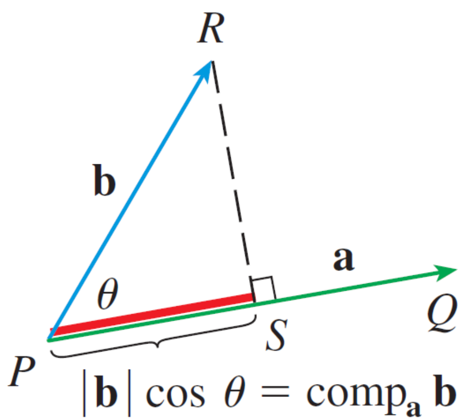




Vector projections

Projections

The scalar projection of **b** onto **a** (also called the **component of b along a**) is defined to be the signed magnitude of the vector projection, which is the number $|\mathbf{b}| \cos \theta$, where θ is the angle between **a** and **b**.
This is denoted by $\text{comp}_{\mathbf{a}}\, \mathbf{b}$ (the component of **b** along **a**).



Scalar projection

Observe that it is negative if $\pi/2 < \theta \leq \pi$

Projections

Scalar projection of **b** onto **a**: $\quad \text{comp}_{\mathbf{a}}\, \mathbf{b} = \dfrac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|}$

Vector projection of **b** onto **a**: $\quad \text{proj}_{\mathbf{a}}\, \mathbf{b} = \left(\dfrac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|}\right) \dfrac{\mathbf{a}}{|\mathbf{a}|} = \dfrac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|^2}\, \mathbf{a}$

Example

Find the scalar projection and vector projection of **b** = [1, 1, 2] onto **a** = [–2, 3, 1].

Solution:

Since $|\mathbf{a}| = \sqrt{(-2)^2 + 3^2 + 1^2} = \sqrt{14}$, the scalar projection of **b** onto **a** is

$$\text{comp}_{\mathbf{a}}\,\mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|}$$

$$= \frac{(-2)(1) + 3(1) + 1(2)}{\sqrt{14}}$$

$$= \frac{3}{\sqrt{14}}$$

Example – *Solution*

cont'd

The vector projection is this scalar projection times the unit vector in the direction of **a**:

$$\text{proj}_{\mathbf{a}}\,\mathbf{b} = \frac{3}{\sqrt{14}}\frac{\mathbf{a}}{|\mathbf{a}|}$$

$$= \frac{3}{14}\mathbf{a}$$

$$= \left[-\frac{3}{7}, \frac{9}{14}, \frac{3}{14}\right]$$

Matrix Notation

A **matrix** is a rectangular array of numbers. We use uppercase symbols to denote matrices.

The **size** of a matrix is defined by the number of rows and columns it contains.

For example, an $m \times n$ matrix has m rows and n columns.

The ij-th entry of A is the entry in the i-th row and the j-th column.
We denote this entry by a_{ij}

Matrix Notation

For example, the matrix

(1) $$A = \begin{bmatrix} 0 & 7 & 1 \\ 2 & 9 & 2 \end{bmatrix}$$

is a 2×3 matrix, and $a_{12} = 7$, $a_{21} = 2$, $a_{23} = 2$, and so forth.

A **square** matrix is one in which the number of rows is the same as the number of columns. We say that a square matrix is $n \times n$, or has size n, if it has n rows and columns.

The **transpose** of a matrix is obtained by interchanging its rows and columns and is denoted by a superscript T

Matrix Notation

For example, A^T is the transpose of the previous matrix A,

(2) $$A^T = \begin{bmatrix} 0 & 2 \\ 7 & 9 \\ 1 & 2 \end{bmatrix}$$

If A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix.

Vectors are often treated using the notation of matrices. All vectors we have seen so far have been written as **row vectors** in that their components are listed as a row.

We can view the row form of a vector as a $1 \times n$ matrix and the column form as an $n \times 1$ matrix.
if $\mathbf{v}$ denotes a vector written in row form, then $\mathbf{v}^T$ is the same vector written in column form.

Tensors

After scalars, vectors, and matrices, we can generalize the idea and obtain a **Tensor**.
A Tensor is a **multidimensional arrays of numbers**.
They are very important in sciences: are used to describe complex physical properties and manipulate multi-dimensional data across fields like physics, engineering, and machine learning.

Types of Tensors by Rank
Rank 0 Tensor (Scalar): A single number with no direction.
Rank 1 Tensor (Vector): A one-dimensional array of numbers, representing a quantity with magnitude and direction.
Rank 2 Tensor (Matrix): A two-dimensional array of numbers.
Rank N Tensor: generalization to N dimensions, used to represent data in an N-dimensional grid.

Sum of two Matrices

Only matrices of the same size can be added. If A and B are both $m \times n$ matrices with entries a_{ij} and b_{ij}, then $A + B$ is a new $m \times n$ matrix whose entries are $a_{ij} + b_{ij}$

Thus the matrix sum $A + B$ is calculated by adding the corresponding entries of each matrix.

Example 1

Evaluate the following sums, if possible.

(a) $M + N$, where $M = \begin{bmatrix} 2 & x & 9 \\ 4 & 5 & 6 \end{bmatrix}$ and $N = \begin{bmatrix} 92 & 6 & 2 \\ 15 & 3 & 1 \end{bmatrix}$

(b) $X + Y$, where $X = \begin{bmatrix} 5 & 3 \\ 7 & 13 \end{bmatrix}$ and $Y = \begin{bmatrix} 3 & 9 & 21 \\ 5 & 7 & 6 \end{bmatrix}$

Solution:

(a) Both matrices are 2×3 and therefore can be added. Adding the corresponding entries gives

$$M + N = \begin{bmatrix} 94 & x + 6 & 11 \\ 19 & 8 & 7 \end{bmatrix}.$$

(b) Matrix X is 2×2 while Y is 2×3. Since they are not the same size, they cannot be added.

Matrix Scalar Multiplication

Scalar multiplication with matrices works just as it does with vectors. If A is an $m \times n$ matrix with entries a_{ij} and c is a scalar, then the product cA is an $m \times n$ matrix with entries ca_{ij}. In other words, the product cA is calculated by multiplying each entry of A by c

Matrix subtraction can be defined through a combination of scalar multiplication and matrix addition. If A and B are both $m \times n$ matrices with entries a_{ij} and b_{ij}, then $A - B$ is calculated by first multiplying B by -1 and then adding this to A

Matrix Scalar Multiplication

Thus the difference $A - B$ is calculated by subtracting entry b_{ij} from a_{ij}. Again notice that matrix subtraction can be performed only with matrices of the same size.

Finally, two matrices A and B are said to be **equal** if $A - B = 0$ where 0 is an $m \times n$ matrix of zeros.

Properties of Matrix Addition If A, B, and C are $m \times n$ matrices and a and b are scalars, then

1. $A + B = B + A$

2. $A + (B + C) = (A + B) + C$

3. $A + 0 = A$

4. $A + (-A) = 0$

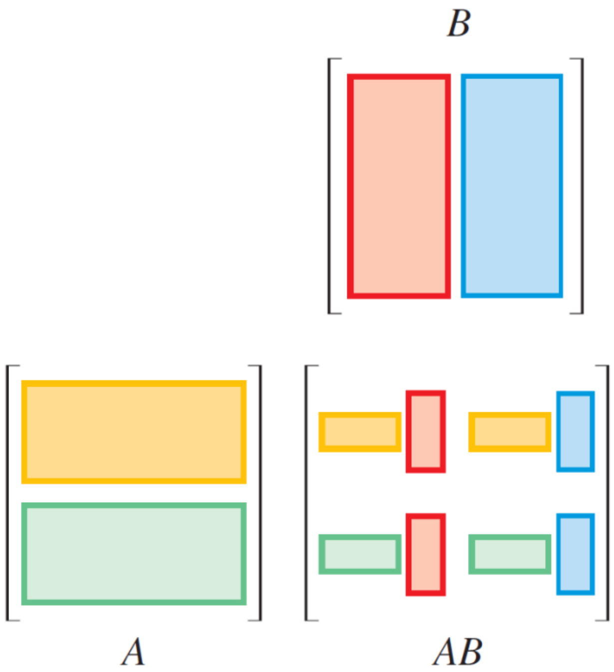
5. $a(A + B) = aA + aB$

6. $(a + b)A = aA + bA$

Matrix Multiplication

If we wish to calculate the matrix product *AB*, we view the matrix *A* as a collection of row vectors and *B* as a collection of column vectors.

The *ij*-th entry of the resulting product is the dot product of the *i*-th row of *A* with the *j*-th column of *B*



Matrix Multiplication

For example, if *A* is a 2 × 3 matrix and *B* is a 3 × 2 matrix, we have

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}$$

Note that the resulting matrix is 2 × 2

Matrix Multiplication

More generally, if A is an $m \times n$ matrix and B is an $n \times p$ matrix, then $C = AB$ is an $m \times p$ matrix whose ij-th entry is the dot product of the i-th row of A with the j-th column of B

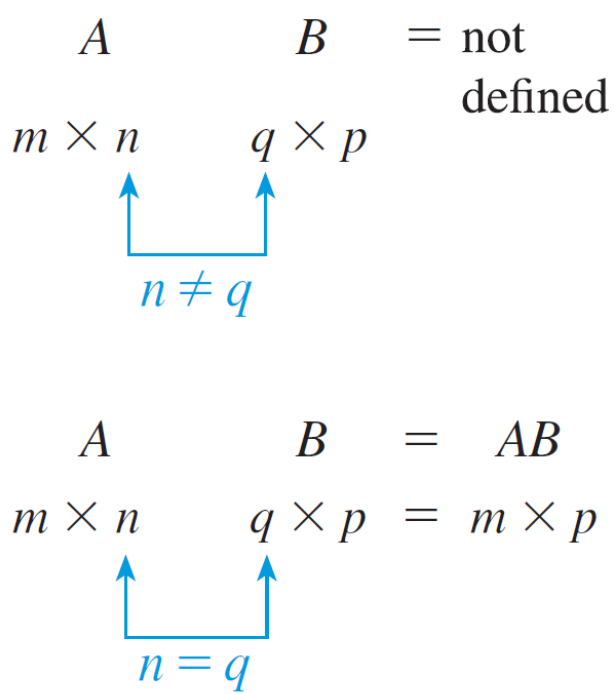
$$\underbrace{\begin{bmatrix} — & \mathbf{r}_1 & — \\ — & \mathbf{r}_2 & — \\ & \vdots & \\ — & \mathbf{r}_m & — \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_p \\ | & | & & | \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{c}_1 & \mathbf{r}_1 \cdot \mathbf{c}_2 & \cdots & \mathbf{r}_1 \cdot \mathbf{c}_p \\ \mathbf{r}_2 \cdot \mathbf{c}_1 & \mathbf{r}_2 \cdot \mathbf{c}_2 & \cdots & \mathbf{r}_2 \cdot \mathbf{c}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_m \cdot \mathbf{c}_1 & \mathbf{r}_m \cdot \mathbf{c}_2 & \cdots & \mathbf{r}_m \cdot \mathbf{c}_p \end{bmatrix}}_{AB}$$

The i-th row of A is indicated by $\boldsymbol{r}_i$
The j-th column of B is indicated by $\boldsymbol{c}_j$

Matrix Multiplication

Matrix multiplication is not defined for matrices where the number of columns of the first matrix is different from the number of rows of the second matrix.

A simple way to determine if a given matrix multiplication is defined is to write the size of the first matrix, followed by the size of the second matrix



Matrix Multiplication

If the two "inner" numbers are not the same, then the matrix multiplication is not possible. If they are the same, then the resulting matrix has size given by the two "outer" numbers.

Matrix multiplication is summarized by the following rule.

Matrix Multiplication If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then their product $C = AB$ is an $m \times p$ matrix whose entries are given by

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant p$.

Example 2

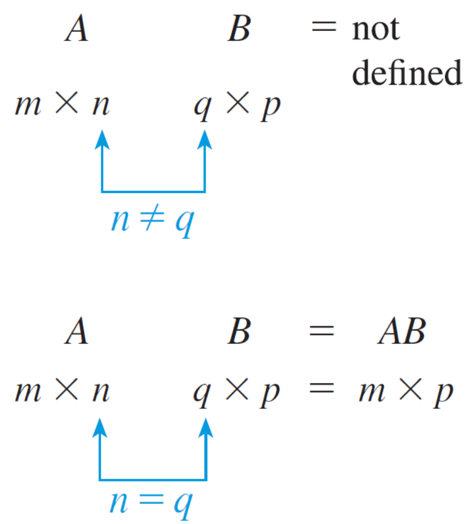
Determine each matrix product if it is defined.

$$A = \begin{bmatrix} 2 & 7 \\ 9 & -3 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & -7 & 2 \\ 1 & 5 & 9 \end{bmatrix} \qquad C = \begin{bmatrix} 5 & -6 \\ 8 & 2 \end{bmatrix}$$

(a) *AB* (b) *BA* (c) *AC* (d) *CA*

Solution:

(a) Matrices *A* and *B* have sizes 2 × 2 and 2 × 3, respectively. From Figure 3, since *n* = *q* = 2, matrix multiplication is therefore defined and the resulting matrix is 2 × 3



Example 2 – *Solution*

cont'd

Performing the calculation, we obtain

$$\begin{bmatrix} 2 & 7 \\ 9 & -3 \end{bmatrix}\begin{bmatrix} 3 & -7 & 2 \\ 1 & 5 & 9 \end{bmatrix} = \begin{bmatrix} 6+7 & -14+35 & 4+63 \\ 27-3 & -63-15 & 18-27 \end{bmatrix} = \begin{bmatrix} 13 & 21 & 67 \\ 24 & -78 & -9 \end{bmatrix}$$

(b) Matrices B and A have sizes 2×3 and 2×2.
Since $n = 3$ and $q = 2$, $n \neq q$ the product is not defined.

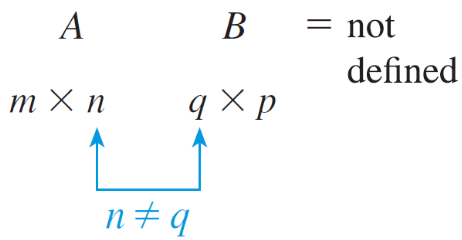
(c) We have $n = q = 2$. The resulting matrix is 2×2.
Performing the calculation, we obtain

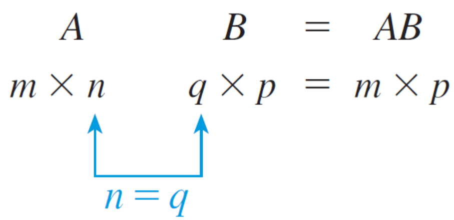
$$\begin{bmatrix} 2 & 7 \\ 9 & -3 \end{bmatrix}\begin{bmatrix} 5 & -6 \\ 8 & 2 \end{bmatrix} = \begin{bmatrix} 10+56 & -12+14 \\ 45-24 & -54-6 \end{bmatrix} = \begin{bmatrix} 66 & 2 \\ 21 & -60 \end{bmatrix}$$

Example 2 – *Solution*

cont'd

(d) We have $n = q = 2$.





The resulting matrix is 2×2 and, performing the calculation, we obtain

$$\begin{bmatrix} 5 & -6 \\ 8 & 2 \end{bmatrix}\begin{bmatrix} 2 & 7 \\ 9 & -3 \end{bmatrix}\begin{bmatrix} 10 - 54 & 35 + 18 \\ 16 + 18 & 56 - 6 \end{bmatrix} = \begin{bmatrix} -44 & 53 \\ 34 & 50 \end{bmatrix}$$

Matrix Multiplication

Multiplication of two quantities α and β is said to be **commutative** if $\alpha\beta = \beta\alpha$. Parts (c) and (d) of Example 2 illustrate the important fact that matrix multiplication is **not commutative** (in general $AC \neq CA$, as in this example).

An $n \times n$ matrix D is called **diagonal** if all off-diagonal entries are zero; that is, $d_{ij} = 0$ for all $i \neq j$

Example 3

For an arbitrary 2×2 diagonal matrix D with entries d_{ii}, calculate the matrix DD

Solution:

Calculating the matrix product, we obtain

$$DD = \begin{bmatrix} d_{11} & 0 \\ 0 & d_{22} \end{bmatrix} \begin{bmatrix} d_{11} & 0 \\ 0 & d_{22} \end{bmatrix} = \begin{bmatrix} d_{11}^2 & 0 \\ 0 & d_{22}^2 \end{bmatrix}$$

Thus DD is a diagonal matrix with entries d_{ii}^2.

Identity Matrix

A diagonal matrix is called an **identity** matrix if all the entries on the diagonal are 1. Identity matrices are usually denoted by I and play the same role in matrix multiplication that the number 1 plays in regular multiplication.

Properties of Matrix Multiplication Suppose A, B, and C are matrices and a and b are scalars. Provided the required matrix multiplications are defined, then

1. $A(BC) = (AB)C$

2. $(aA)(bB) = abAB$

3. $A(B + C) = AB + AC$

4. $(B + C)A = BA + CA$

5. $IA = A, AI = A$

6. $0A = 0, A0 = 0$

Note: Matrix multiplication is *not*, in general, commutative; that is, $AB \neq BA$.

The Inverse of a Matrix

Note that $AI = A$ and $IA = A$ just as $a1 = a$ and $1a = a$ in scalar multiplication.
Given a scalar a we can define a^{-1} such that $a^{-1}a = 1$, where $a^{-1} = 1/a$ provided that $a \neq 0$.

In other words, a^{-1} , which is called the inverse of *a,* is the quantity that, when multiplied with a, gives 1. And if $a = 0$, then no such quantity exists.

Also for a matrix A we can write an inverse matrix $B = A^{-1}$

The Inverse of a Matrix

Definition Suppose that A is an $n \times n$ matrix. If there exists an $n \times n$ matrix B such that

$$AB = BA = I$$

then B is called the **inverse** of A and is denoted by A^{-1}.

Note that, when the inverse of a matrix exists, it is unique. Also, if there exists a matrix B such that $AB = I$, then necessarily $BA = I$ as well (and vice versa).

Therefore we need check only one order of multiplication when finding an inverse. If A has an inverse, then we say that A is **invertible** or **nonsingular**. Otherwise A is called **singular**.

Example 1

Show that *B* is the inverse of *A*, where

$$A = \begin{bmatrix} 1 & 2 \\ 7 & 5 \end{bmatrix} \qquad B = \begin{bmatrix} -\frac{5}{9} & \frac{2}{9} \\ \frac{7}{9} & -\frac{1}{9} \end{bmatrix}$$

Solution:

Calculating the matrix product gives

$$BA = \begin{bmatrix} -\frac{5}{9} & \frac{2}{9} \\ \frac{7}{9} & -\frac{1}{9} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 7 & 5 \end{bmatrix} = \begin{bmatrix} -\frac{5}{9}(1) + \frac{2}{9}(7) & -\frac{5}{9}(2) + \frac{2}{9}(5) \\ \frac{7}{9}(1) - \frac{1}{9}(7) & \frac{7}{9}(2) - \frac{1}{9}(5) \end{bmatrix}$$

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

$$= I$$

Example 2

Derive the inverse of the matrix *A* from Example 1 by using the definition of an inverse.

Solution:

From the definition, if a matrix $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ is the inverse of $A = \begin{bmatrix} 1 & 2 \\ 7 & 5 \end{bmatrix}$, then *AB* = *BA* = *I*. We need focus on only one of these orders of mulitplication.

Choosing *AB* = *I*, we have

$$\begin{bmatrix} 1 & 2 \\ 7 & 5 \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} b_{11} + 2b_{21} & b_{12} + 2b_{22} \\ 7b_{11} + 5b_{21} & 7b_{12} + 5b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Example 2 – *Solution*

cont'd

Looking at this equation entry by entry, we see that it represents four equations in four unknowns.

The four equations can be split into two pairs of equations, each with two unknowns:

$$\begin{aligned} b_{11} + 2b_{21} &= 1 \\ 7b_{11} + 5b_{21} &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} b_{12} + 2b_{22} &= 0 \\ 7b_{12} + 5b_{22} &= 1 \end{aligned}$$

We can now solve each pair by substitution. If we focus on the first pair, we see that the second equation gives $b_{21} = -7b_{11}/5$.

Example 2 – *Solution*

cont'd

Substituting this into the first equation of this pair then gives $b_{11} + 2(-7b_{11}/5) = 1$, whose solution is $b_{11} = -5/9$. This can then be back-substituted into $b_{21} = -7b_{11}/5$ to give $b_{21} = 7/9$.

Similarly, the first equation of the second pair gives $b_{12} = -2b_{22}$. Substituting this into the second equation of this pair then gives $7(-2b_{22}) + 5b_{22} = 1$, whose solution is $b_{22} = -1/9$.

This can then be back-substituted into $b_{12} = -2b_{22}$ to give $b_{12} = 2/9$. Putting these results together in the matrix B gives

$$B = \begin{bmatrix} -\frac{5}{9} & \frac{2}{9} \\ \frac{7}{9} & -\frac{1}{9} \end{bmatrix}$$

The Inverse of a Matrix

The Inverse of a 2 x 2 Matrix Suppose

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Then A is invertible and

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

If $a_{11}a_{22} - a_{12}a_{21} = 0$, then A is not invertible (that is, A is singular).

Example 3

If possible, find the inverse.

(a) $M = \begin{bmatrix} 7 & 9 \\ 5 & 6 \end{bmatrix}$ (b) $N = \begin{bmatrix} 14 & 6 \\ 7 & 3 \end{bmatrix}$ (c) $J = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$

Solution:

From the formula for the inverse of a 2 × 2 matrix, an inverse will exist if and only if $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

(a) $m_{11}m_{22} - m_{12}m_{21} = (7)(6) - (9)(5) = 42 - 45 = -3$

Therefore an inverse exists. Using the formula, we obtain

$$M^{-1} = \frac{1}{-3}\begin{bmatrix} 6 & -9 \\ -5 & 7 \end{bmatrix} = \begin{bmatrix} -2 & 3 \\ \frac{5}{3} & -\frac{7}{3} \end{bmatrix}$$

Example 3 – *Solution*

cont'd

(b) $n_{11}n_{22} - n_{12}n_{21}$ = (14)(3) – (6)(7) = 42 – 42 = 0

Therefore an inverse does not exist.

(c) $j_{11}\, j_{22} - j_{12}\, j_{21}$ = (2)(–2) – (3)(–1) = –4 + 3 = –1

Therefore an inverse exists.

We find it using the formula. Note that this last example has the interesting property that $J^{-1} = J$. In other words, J is its own inverse.

$$J^{-1} = \frac{1}{-1}\begin{bmatrix} -2 & -3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$$

The Inverse of a Matrix

Properties of Matrix Inverses Suppose A and B are both invertible $n \times n$ matrices. Then

1. $(A^{-1})^{-1} = A$

2. $(AB)^{-1} = B^{-1}A^{-1}$

3. A^{-1} is unique.

The Determinant of a Matrix

To any $n \times n$ matrix A we can assign a scalar quantity called its *determinant*, denoted by det A. If we view scalars as 1×1 matrices, then the definition of the determinant for matrices of sizes $n = 1$ through $n = 3$ is as follows.

The Determinant Suppose A is an $n \times n$ matrix.

1. If $n = 1$, then $\det A = a_{11}$.
2. If $n = 2$, then $\det A = a_{11}a_{22} - a_{12}a_{21}$.
3. If $n = 3$, then $\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$.

The Determinant of a Matrix

Given the determinant of a matrix we have the following theorem:

> **(1) Theorem** If A is an $n \times n$ matrix, then A is invertible if and only if $\det A \neq 0$.

Notice that the quantity in the denominator of the formula for the inverse of a 2×2 matrix is its determinant.

Example 4

Which of the following matrices are invertible?

(a) $A = \begin{bmatrix} 2 \end{bmatrix}$ (b) $B = \begin{bmatrix} 2 & 3 \\ 6 & 9 \end{bmatrix}$ (c) $C = \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix}$ (d) $D = \begin{bmatrix} 10 & 7 & 3 \\ 13 & 5 & 8 \\ 6 & -1 & 7 \end{bmatrix}$

Solution:

(a) The matrix A is 1×1, with det $A = 2$. Therefore it is invertible.

(b) The matrix B is 2×2, with det $B = (2)(9) - (3)(6) = 0$. Therefore it is not invertible.

(c) The matrix C is 2×2, with det $C = (5)(1) - (3)(2) = -1$. Therefore it is invertible.

Example 4 – *Solution*

cont'd

(d) The matrix D is 3×3, and

$$\begin{aligned}\det D &= (10)(5)(7) + (7)(8)(6) + (3)(13)(-1) - (3)(5)(6) - (10)(8)(-1) - (7)(13)(7) \\ &= 350 + 336 + (-39) - 90 - (-80) - 637 \\ &= 0\end{aligned}$$

Therefore D is not invertible.

Solving Systems of Linear Equations

Matrix inverses (and therefore determinants) are very useful for solving linear systems of *n* equations with *n* unknowns. Consider, for example, the following pair of equations with two unknowns:

$$3x_1 - 2x_2 = -4$$

$$7x_1 + x_2 = 19$$

Writing these equations in matrix notation gives

$$A\mathbf{x} = \mathbf{b} \tag{2}$$

With $A = \begin{bmatrix} 3 & -2 \\ 7 & 1 \end{bmatrix}$ matrix of coefficients, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} -4 \\ 19 \end{bmatrix}$.

Solving Systems of Linear Equations

Such a system of equations in which $\mathbf{b} \neq \mathbf{0}$ is called an **non-homogeneous** (or inhomogeneous) system.

In general, there are three possibilities for such a system:
(i) there is a unique nonzero solution for $\mathbf{x}$,
(ii) there are infinitely many solutions for $\mathbf{x}$, or
(iii) there is no solution for $\mathbf{x}$.

If, instead of a system, we had only a scalar equation,

$$ax = b$$

Provided that $a \neq 0$ we can solve it by multiplying both sides by $a^{-1} = 1/a$ to get $a^{-1}ax = a^{-1}b$, or $x = a^{-1}b$

Solving Systems of Linear Equations

In a similar fashion, if $\det A \neq 0$ the matrix A will be invertible and we can multiply both sides of the system by A^{-1} to obtain

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$$

or

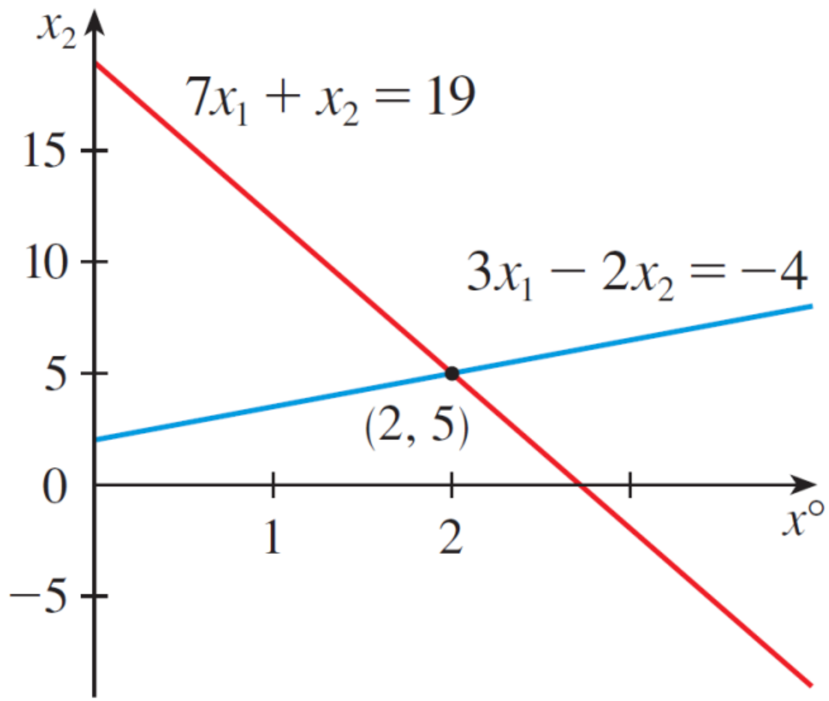
$$\mathbf{x} = A^{-1}\mathbf{b}$$

Using the formula for the inverse of a 2×2 matrix, we obtain

$$\mathbf{x} = \frac{1}{(3)(1) - (-2)(7)}\begin{bmatrix} 1 & 2 \\ -7 & 3 \end{bmatrix}\begin{bmatrix} -4 \\ 19 \end{bmatrix} = \begin{bmatrix} \frac{1}{17} & \frac{2}{17} \\ -\frac{7}{17} & \frac{3}{17} \end{bmatrix}\begin{bmatrix} -4 \\ 19 \end{bmatrix} = \begin{bmatrix} \dfrac{-4 + 38}{17} \\ \dfrac{28 + 57}{17} \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

Solving Systems of Linear Equations

That system thus has a unique solution given by $x_1 = 2$, $x_2 = 5$. Graphically, the two equations represent straight lines in the x_1x_2-plane and their intersection point is the unique solution



Solving Systems of Linear Equations

More generally, we have the following theorem.

(4) Theorem Suppose A is an $n \times n$ matrix, $\mathbf{x}$ is an $n \times 1$ vector of unknowns, and $\mathbf{b}$ is an $n \times 1$ vector of constants. If A is invertible, then the inhomogeneous system of equations $A\mathbf{x} = \mathbf{b}$ has a unique solution given by $\mathbf{x} = A^{-1}\mathbf{b}$.

If A in Theorem 4 is not invertible, then the system of equations can have either infinitely many solutions or no solution.

Example 5

Using $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, obtain the solution to the inhomogeneous system of equations for each of the following matrices.

(a) $A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ (c) $C = \begin{bmatrix} 1 & -1 \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$

Solution:

(a) First we calculate det A = 1 – (–2) = 3 and therefore the matrix A is invertible.

Theorem 4 then tells us that the inhomogeneous equation $A\mathbf{x} = \mathbf{b}$ has a unique solution given by $\mathbf{x} = A^{-1}\mathbf{b}$.

Example 5 – *Solution*

cont'd

Using the formula of the inverse of a 2 × 2 matrix, we obtain

$$\mathbf{x} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

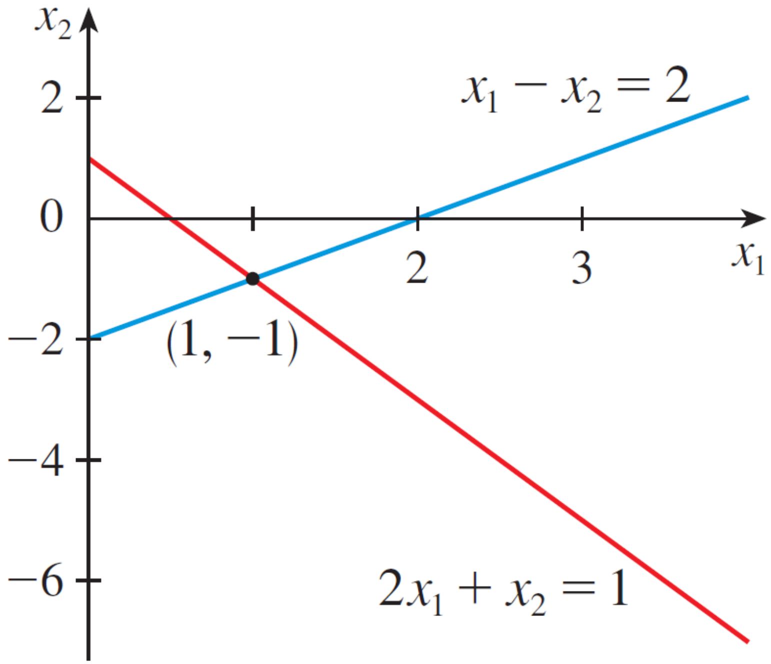
$$= \begin{bmatrix} \dfrac{2 + 1}{3} \\ \dfrac{-4 + 1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Example 5 – *Solution*

cont'd

The solution corresponds to the intersection point of the lines defined by $x_1 - x_2 = 2$ and $2x_1 + x_2 = 1$



Example 5 – *Solution*

cont'd

(b) Calculating the determinant of B gives det $B = 1 – 1 = 0$. Therefore the matrix B is not invertible. Thus, the inhomogeneous equation $B\mathbf{x} = \mathbf{b}$ might have an infinite number of solutions or no solution.

To determine which is the case, we need to investigate the equations in more depth. Carrying out the matrix multiplication gives the two equations $x_1 – x_2 = 2$ and $–x_1 + x_2 = 1$

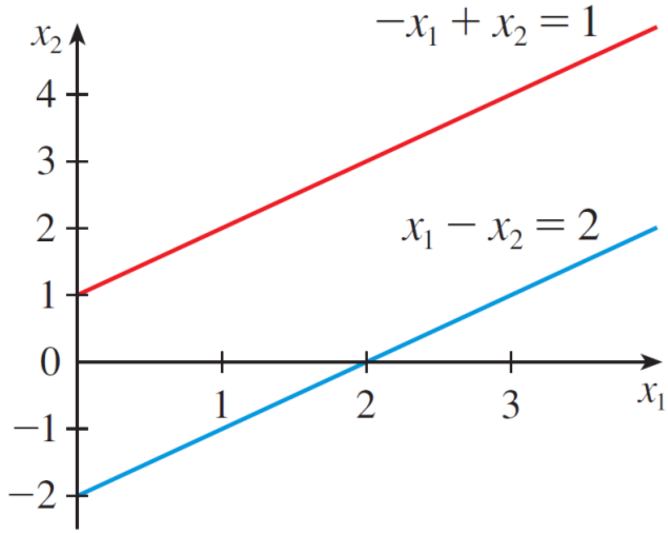
Solving the second equation for x2 gives x2 = 1 + x1, and substituting this into the first equation gives
x1 – (1 + x1) = 2

Example 5 – *Solution*

cont'd

This simplifies to $-1 = 2$. Because there is no choice of x_1 that will make this true, there is no solution.

Graphically, the two equations represent parallel straight lines in the x_1x_2-plane and the lack of a solution corresponds to their not having an intersection point.



Example 5 – *Solution*

cont'd

(c) Calculating the determinant of C gives det $C = -\frac{1}{2} - \left(-\frac{1}{2}\right) = 0.$

Therefore the matrix C is not invertible and the inhomogeneous equation $C\mathbf{x} = \mathbf{b}$ might have an infinite number of solutions or no solution.

Carrying out the matrix multiplication gives the two equations $x_1 - x_2 = 2$ and $\frac{1}{2}x_1 - \frac{1}{2}x_2 = 1.$

Solving the second equation for x_1 gives $x_1 = 2 + x_2$, and substituting this into the first equation gives

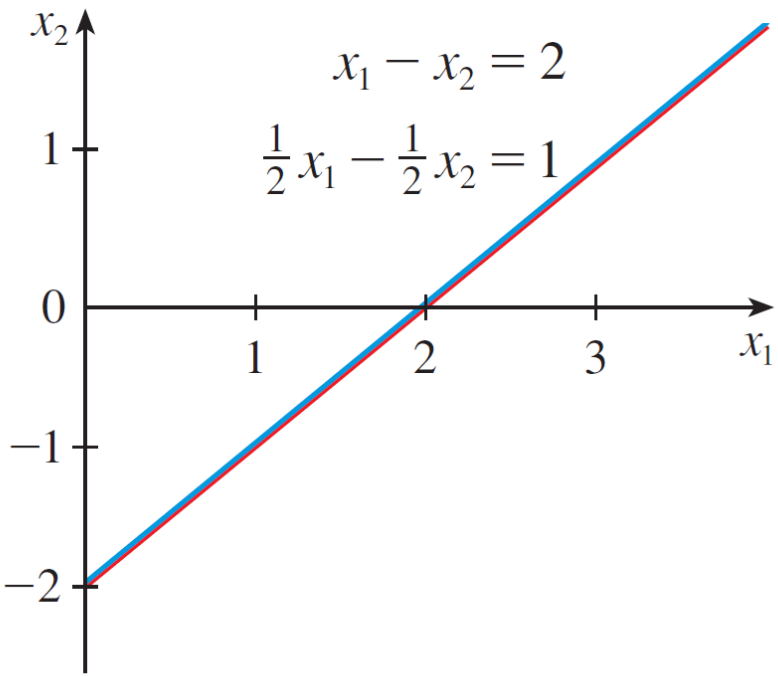
$(2 + x_2) – x_2 = 2.$

This simplifies to 2 = 2. This equation holds true no matter what value of x_2 is chosen and therefore there are an infinite number of solutions.

Example 5 – *Solution*

cont'd

Graphically, the two equations represent the same straight line in the x_1x_2-plane, and therefore there are an infinite number of points of intersection



Solving Systems of Linear Equations

The solution to a system of linear equations is simplified considerably in the special case where $\mathbf{b} = \mathbf{0}$. This results in an equation of the form $A\mathbf{x} = \mathbf{0}$, which is called a **homogeneous** system.

Clearly $\mathbf{x} = \mathbf{0}$ is always a solution and it is referred to as the **trivial solution**. There are therefore now only two possibilties: (i) the trivial solution is the unique solution, or (ii) there are an infinite number of nontrivial solutions.

Solving Systems of Linear Equations

The following theorem tells us when each of these outcomes occurs.

(5) Theorem Suppose A is an $n \times n$ matrix, $\mathbf{x}$ is an $n \times 1$ vector of unknowns, and $\mathbf{0}$ is an $n \times 1$ vector of zeros. If A is invertible, then the homogeneous system of equations $A\mathbf{x} = \mathbf{0}$ has a unique solution given by the trivial solution $\mathbf{x} = \mathbf{0}$. If A is not invertible, then there are infinitely many nontrivial solutions.

Connection with Cramer's Rule

- **System of equations** $\begin{cases} 2x + y = 5 \\ x + 3y = 7 \end{cases}$ where $A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, $b = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$

Method using the inverse

- Determinant: det(A) = 2·3 −1·1 = 6−1 = 5
- Adjugate matrix: adj$(A) = \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$ Inverse: A−1 = 1/det(A) adj(A) = 1/5 $\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$
- $x = A^{-1}b = \frac{1}{5}\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ 7 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 3 \cdot 5 + (-1) \cdot 7 \\ -1 \cdot 5 + 2 \cdot 7 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 15 - 7 \\ -5 + 14 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 8 \\ 9 \end{bmatrix}$
- Hence $x = \frac{1}{5}\begin{bmatrix} 8 \\ 9 \end{bmatrix} = \begin{bmatrix} \frac{8}{5} \\ \frac{9}{5} \end{bmatrix}$

Cramer's Rule

- For x_1: in A replace the first column with b
- A1= $\begin{bmatrix} 5 & 1 \\ 7 & 3 \end{bmatrix}$, det(A1) = 5·3 − 1·7 = 15−7 = 8 → x_1 = 8/5
- For x_2: in A replace the second column with b
- A2= $\begin{bmatrix} 2 & 5 \\ 1 & 7 \end{bmatrix}$, det(A2) = 2·7 − 5·1 = 14−5 = 9 → x_2 = 9/5

So, both methods give the same solution: $x = \left(\frac{8}{5}, \frac{9}{5}\right)$
Cramer's rule is essentially a consequence of the formula of the inverse, since the inverse can be expressed in terms of determinants and cofactors.

How Matrix Multiplication Changes Vectors?

In the earlier section we studied models for the dynamics of vectors having the form

$$\mathbf{n}_{t+1} = A\mathbf{n}_t$$

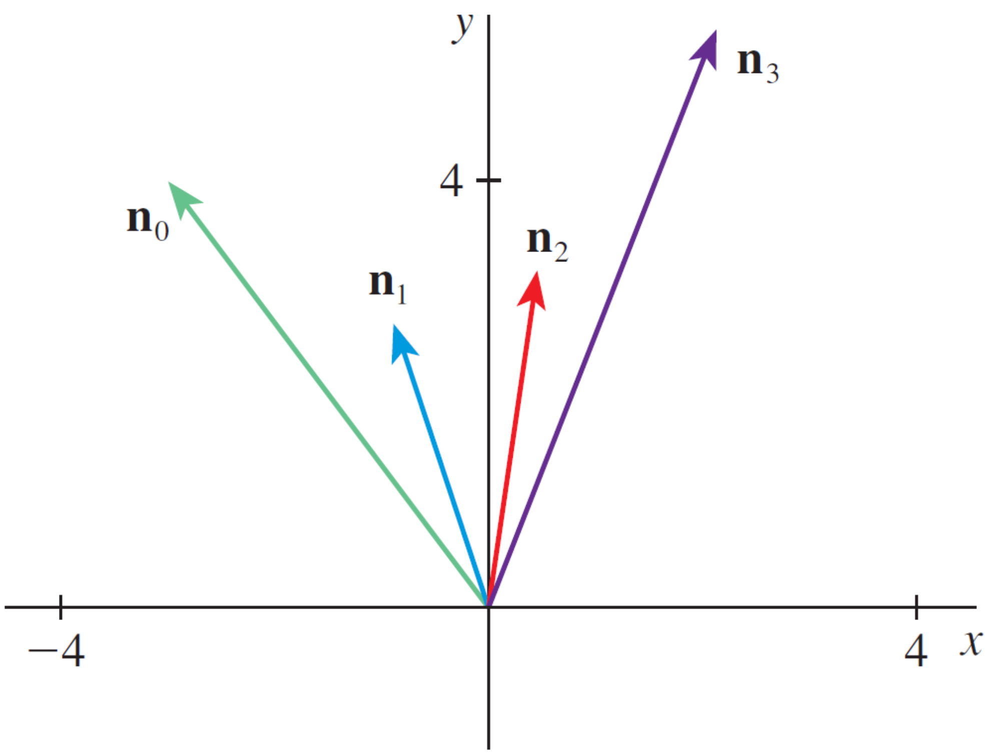
where $\mathbf{n}_t$ is a vector of variables and A is a square matrix. Our goal now is to characterize in general what such multiplication of a vector by a matrix does to the vector.

As an example, suppose that

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

How Matrix Multiplication Changes Vectors?

The matrix multiplication changes the vector from one time step to the next when $\mathbf{n}_0 = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$.



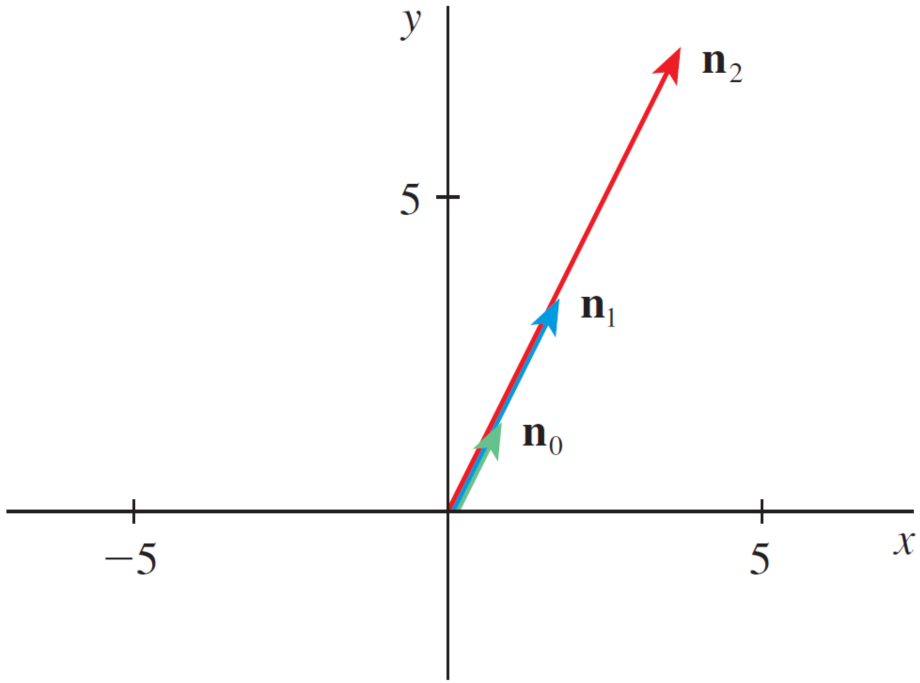
How Matrix Multiplication Changes Vectors?

The changes in the vector look quite complicated. For example, the direction of the vector keeps changing and sometimes the vector is compressed from one step to the next (for example, between time 0 and time 1) and sometimes it is stretched (for example, between time 2 and time 3).

If you were to experiment with different initial vectors $\mathbf{n}_0$, you would get different patterns. However, you might happen across an initial vector that results in a particularly simple result.

How Matrix Multiplication Changes Vectors?

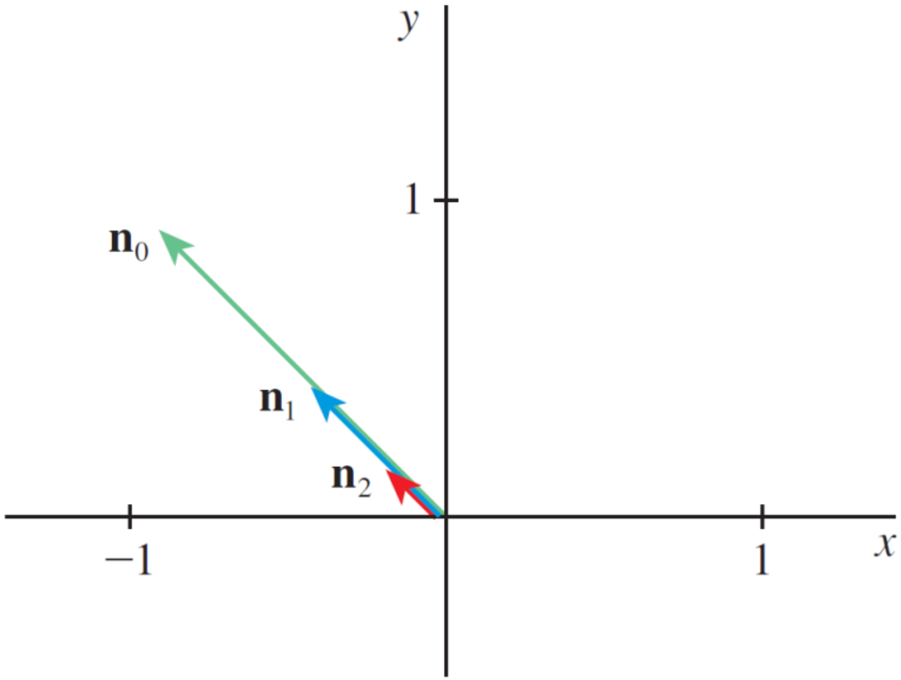
For example, consider $\mathbf{n}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. In this case the vector remains on its initial axis, and it is simply stretched by a factor of 2 each time step, as shown in the figure.



How Matrix Multiplication Changes Vectors?

Likewise, if $\mathbf{n}_0 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ then again the vector remains on its initial axis, but now it is compressed by a factor of 1/2 each time step, as shown in the figure.

Thus, for some special initial vectors, the change in the vector is relatively easy to describe.



Eigenvectors and Eigenvalues

An eigenvector of a matrix is a vector that, when multiplied by the matrix, is simply changed in length.

(1) Definition Suppose that A is an $n \times n$ matrix. A *nonzero* vector $\mathbf{v}$ that satisfies the equation

$$A\mathbf{v} = \lambda\mathbf{v}$$

is called an **eigenvector** of the matrix A. The scalar λ is the **eigenvalue** associated with this eigenvector.

How do we find eigenvectors and eigenvalues? As the following calculations reveal, it is easiest to first calculate the eigenvalues of a matrix and then calculate their associated eigenvectors.

Eigenvectors and Eigenvalues

Consider the matrix $A = \begin{bmatrix} 1 & \frac{1}{2} \\ 1 & \frac{3}{2} \end{bmatrix}$. From Definition (1) we have

$$A\mathbf{v} = \lambda\mathbf{v}$$

or, equivalently,

$$A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}$$

where **0** is the zero vector.

Eigenvectors and Eigenvalues

To proceed further, we factor out the vector **v**. To do so, we must first multiply λ by the identity matrix I in order for the elements of the equation to maintain compatible sizes:

(2)

$$(A - \lambda I)\mathbf{v} = \mathbf{0}$$

Notice that, although λ is a scalar, $A - \lambda I$ is a 2×2 matrix.

We now seek a vector **v** that satisfies Equation 2. We know that if the matrix $A - \lambda I$ is invertible, then the only solution is the trivial solution $\mathbf{v} = \mathbf{0}$.

Eigenvectors and Eigenvalues

Therefore, to have a nonzero eigenvector **v**, we require that $A - \lambda I$ be singular. This requires that

$$\det(A - \lambda I) = 0$$

Our considerations have allowed us to remove **v** from the equation, and therefore we now have an equation that determines the eigenvalues λ. Calculating the matrix $A - \lambda I$, we get

$$A - \lambda I = \begin{bmatrix} 1 - \lambda & \frac{1}{2} \\ 1 & \frac{3}{2} - \lambda \end{bmatrix}$$

and, using the definition of the determinant of a 2×2 matrix, we have

$$\det \begin{bmatrix} 1 - \lambda & \frac{1}{2} \\ 1 & \frac{3}{2} - \lambda \end{bmatrix} = (1 - \lambda)\left(\frac{3}{2} - \lambda\right) - \frac{1}{2} \cdot 1$$

Eigenvectors and Eigenvalues

$$= \frac{3}{2} - \lambda - \frac{3}{2}\lambda + \lambda^2 - \frac{1}{2}$$

$$= \lambda^2 - \frac{5}{2}\lambda + 1$$

Setting this result equal to zero and multiplying the equation by 2 then gives $2\lambda^2 - 5\lambda + 2 = 0.$

This can be factored to produce $(\lambda - 2)(2\lambda - 1) = 0,$ the solutions of which are $\lambda = 2$ and $\lambda = ½$. These are the eigenvalues of *A*

Eigenvectors and Eigenvalues

We can now calculate the eigenvector associated with each of the eigenvalues $\lambda = 2$ and $\lambda = \frac{1}{2}$. First let's find the eigenvector associated with $\lambda = 2$.

We seek a vector **v** such that, when $\lambda = 2$ is substituted into the left side of Equation 2, we get the zero vector.

In other words,

$$\begin{bmatrix} 1-\lambda & \frac{1}{2} \\ 1 & \frac{3}{2}-\lambda \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 1-2 & \frac{1}{2} \\ 1 & \frac{3}{2}-2 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Eigenvectors and Eigenvalues

This gives the following pair of equations in two unknowns:

$$-v_1 + \frac{1}{2}v_2 = 0$$

$$v_1 - \frac{1}{2}v_2 = 0$$

These two equations are redundant because both specify that $2v_1 = v_2$. As a result, there are infinitely many solutions—we are free to choose either v_1 or v_2 arbitrarily, and the other is then determined by this choice.

To make our calculations simple, it is usually best to work with whole numbers. For example, we might choose $v_1 = 1$, in which case we then have $v_2 = 2$.

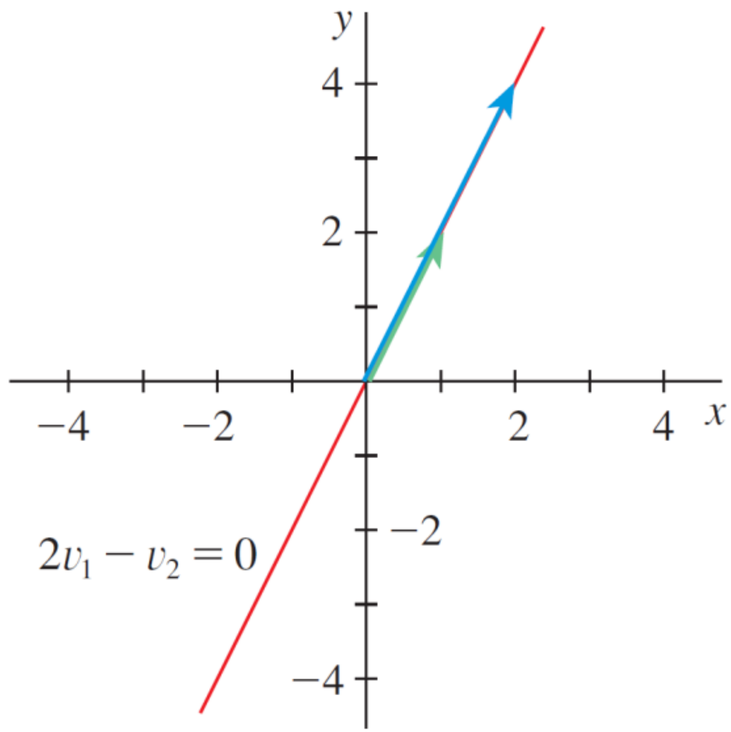
Eigenvectors and Eigenvalues

The eigenvector associated with eigenvalue $\lambda = 2$ is then $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. If we had made a different choice for v_1, we would have ended up with a different vector **v**. For example, if we choose $v_1 = 2$, we would then have $v_2 = 4$ and thus $\mathbf{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$.

Eigenvectors and Eigenvalues

Regardless of our choice, however, all the resulting vectors are eigenvectors and all lie on the same line because they all have the form $\mathbf{v} = \begin{bmatrix} a \\ 2a \end{bmatrix}$ or $\mathbf{v} = a\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ for some scalar *a*

All eigenvectors associated with the eigenvalue λ = 2 lie on the red line.



Eigenvectors and Eigenvalues

We now calculate the eigenvector associated with $\lambda = \frac{1}{2}$.
We obtain

$$\begin{bmatrix} 1-\lambda & \frac{1}{2} \\ 1 & \frac{3}{2}-\lambda \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 1-\frac{1}{2} & \frac{1}{2} \\ 1 & \frac{3}{2}-\frac{1}{2} \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which gives the pair of equations

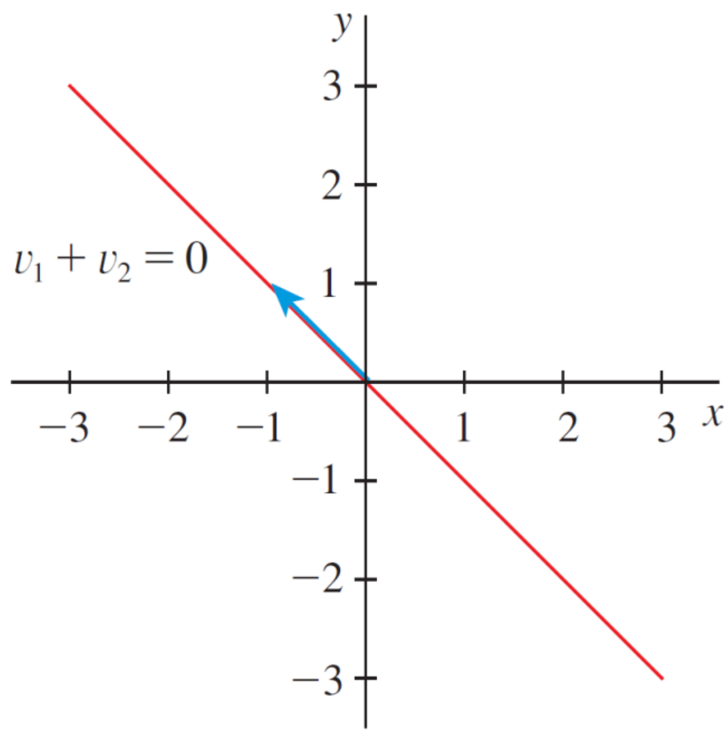
$$\begin{aligned} \tfrac{1}{2}v_1 + \tfrac{1}{2}v_2 &= 0 \\ v_1 + v_2 &= 0 \end{aligned}$$

Again, these equations are redundant, both specifying that $v_1 = -v_2$

Eigenvectors and Eigenvalues

Choosing v_2 = 1, we then have v_1 = –1 and therefore $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$

Again the choice v_2 = 1 is arbitrary, but all choices result in vectors lying on the same line because they all have the form $\mathbf{v} = a\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ for some scalar *a*



All eigenvectors associated with the eigenvalue λ = 1/2 lie on the red line.

Eigenvectors and Eigenvalues

For a square matrix A of size n, the equation $\det(A - \lambda I) = 0$ that determines its eigenvalues is an nth-degree polynomial in λ. This polynomial is referred to as the **characteristic polynomial** of the matrix A.

Given that the eigenvalues of a matrix are the roots of a polynomial, we also must expect that eigenvalues are sometimes complex numbers.

In general, if we define the quantity $i = \sqrt{-1}$, the eigenvalues of matrices composed of real numbers always come in complex conjugate pairs, having the form $\lambda = a + bi$ and $\lambda = a - bi$ for some real numbers a and b

Example 3

Consider the matrix $A = \begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$.

(a) Find its eigenvalues.

(b) Find the eigenvectors associated with the eigenvalues from part (a).

(c) Consider an arbitrary initial vector $\mathbf{n}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$. How does multiplication of $\mathbf{n}_0$ by A affect the length of this vector?

(d) Use the dot product to determine how multiplication by A affects the direction of the vector.

(e) Describe, overall, what multiplication by the matrix A does to vectors.

Example 3(a) – *Solution*

The characteristic polynomial is $(\sqrt{3} - \lambda)(\sqrt{3} - \lambda) - (-1) = 0$
$\rightarrow \lambda^2 - 2\sqrt{3}\lambda + 3 + 1 = 0$

we solve it using the second degree eq. formula

$$\frac{-b \pm \sqrt{b2-4ac}}{2a}$$

to find the two (complex in this case) eigenvalues:

$$\lambda = \tfrac{1}{2}\left[2\sqrt{3} \pm \sqrt{\left(-2\sqrt{3}\right)^2 - 4(4)}\right]$$

$$= \tfrac{1}{2}\left(2\sqrt{3} \pm \sqrt{12 - 16}\right)$$

$$= \sqrt{3} \pm i$$

Example 3(b) – *Solution*

Beginning with eigenvalue $\lambda = \sqrt{3} + i$, we have

$$\begin{bmatrix} \sqrt{3} - \lambda & -1 \\ 1 & \sqrt{3} - \lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Both of these equations specify that $v_2 = -iv_1$. We choose $v_1 = 1$, giving $\mathbf{v} = \begin{bmatrix} 1 \\ -i \end{bmatrix}$.

Example 3(b) – *Solution*

For eigenvalue $\lambda = \sqrt{3} - i,$ we have

$$\begin{bmatrix} \sqrt{3} - \lambda & -1 \\ 1 & \sqrt{3} - \lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

These equations both specify that $v_2 = iv_1$. We choose $v_1 = 1$, giving $\mathbf{v} = \begin{bmatrix} 1 \\ i \end{bmatrix}$. Notice that, when the eigenvalues are complex, their corresponding eigenvectors are also complex.

Example 3(c) – *Solution*

Carrying out the matrix multiplication with the generic vector

$\mathbf{n}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ gives

$$\mathbf{n}_1 = \begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} \sqrt{3}\,x_0 - y_0 \\ x_0 + \sqrt{3}\,y_0 \end{bmatrix}$$

The length of the initial vector is $|\mathbf{n}_0| = \sqrt{x_0^2 + y_0^2}$, and the length of the vector $\mathbf{n}_1$ is

$$|\mathbf{n}_1| = \sqrt{\left(\sqrt{3}\,x_0 - y_0\right)^2 + \left(x_0 + \sqrt{3}\,y_0\right)^2}$$

Example 3(c) – *Solution*

$$= \sqrt{3x_0^2 - 2\sqrt{3}x_0y_0 + y_0^2 + x_0^2 + 2\sqrt{3}x_0y_0 + 3y_0^2}$$

$$= 2\sqrt{x_0^2 + y_0^2}$$

Therefore multiplication by A increases the length of the vector by a factor of 2.

Example 3(d) – *Solution*

From the definition of the dot product, we have

$$\cos\theta = \frac{\mathbf{n}_0 \cdot \mathbf{n}_1}{|\mathbf{n}_0||\mathbf{n}_1|}$$

where θ is the angle between $\mathbf{n}_0$ and $\mathbf{n}_1$. Substituting the vectors $\mathbf{n}_0$ and $\mathbf{n}_1$ gives

$$\cos\theta = \frac{[x_0, y_0] \cdot \left[\sqrt{3}x_0 - y_0, x_0 + \sqrt{3}y_0\right]}{2\sqrt{x_0^2 + y_0^2}\sqrt{x_0^2 + y_0^2}}$$

$$= \frac{\sqrt{3}x_0^2 - x_0y_0 + x_0y_0 + \sqrt{3}\,y_0^2}{2(x_0^2 + y_0^2)} \quad = \frac{\sqrt{3}}{2}$$

Example 3(d) – *Solution*

Solving for θ shows that the angle between n0 and n1 is $\theta = \pi/6$. Thus multiplication by A rotates the vector by 30 degrees.

Example 3(e) – *Solution*

From parts (c) and (d) and from the figure, we see that each multiplication by A rotates the vector 30 degrees in the counterclockwise direction and stretches its length by a factor of 2.