Engineering Mathematics 1

Dr Tang Wee Kee
Division of Mathematical Sciences,
School of Physical and Mathematical Sciences,
Nanyang Technological University,
Singapore

Semester 1 2021/22

Contents

Part I

Algebra

Chapter 1

Complex Numbers

1.1 The set of real numbers $\mathbb{R}$ & its subsets

The following are some commonly used sets of numbers.

1. The set of **natural numbers**, $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

2. The set of **integers** $\mathbb{Z} = \{\ldots, -3, -2, -2, -1, 0, 1, 2, 3, \ldots\}$.

3. The set of **rational numbers** $\mathbb{Q} = \left\{\frac{p}{q} : p, q \in \mathbb{Z} \text{ and } q \neq 0\right\}$.

   A number is rational if and only if it has a recurring decimal representation. (E.g., $0.111.... = \frac{1}{9}$, $\frac{1}{2} = 0.5 = 0.500000....$).

   Question: Is there a rational number x such that $x^2 = 2$?

   Answer: No. The set $\mathbb{Q}$ of rational numbers is "incomplete". It has many holes.

   Numbers with non-recurring decimal representations are called irrational numbers. E.g., $\sqrt{2}, \pi, e$ are irrationals. The sets of rational numbers and irrational numbers together form the set of real numbers.

4. The set of **real** numbers is denoted by $\mathbb{R}$.

   Question: Is the set of real numbers sufficient to handle all our mathematical encounters?

1.2 Complex Numbers

Does the quadratic equation $x^2 + 1 = 0$ have a real root? That is, are there real numbers x at which $x^2 = -1$?

The answer is NO, because we have $x^2 \geq 0$ for every real number x.

To deal with the above irreducible quadratic equation, a new symbol 'i' is introduced, where $i^2 = -1$. Thus, $x^2 + 1 = 0$ has two distinct roots namely i and $-i$.

Powers of i

$$i^2 = -1, \quad i^3 = (i^2)(i) = -i,$$

$$i^4 = (i^2)(i^2) = (-1)(-1) = 1, \quad i^5 = (i^4)(i) = i, \ldots$$

Let $k \in \mathbb{Z}$. Then we have

$$i^{4k} = (i^4)^k = 1, i^{4k+1} = i, i^{4k+2} = -1, i^{4k+3} = -i.$$

Note: Values of i^n depends on the remainder when n is divided by 4.

Definition 1.2.1. *A complex number z is a mathematical object of the form $x + yi$, where x and y are real numbers.*

Remark 1.2.2. *The real numbers x and y are called the real part and imaginary part of the complex number z respectively. We denote the real and imaginary parts of a complex number z by $Re(z)$ and $Im(z)$ respectively.*

We represent the set of all complex numbers by $\mathbb{C}$.

Example 1.2.3. *Examples of complex numbers are:*

$$3 + 5i, \quad 3.5 - i, \quad -\sqrt{3} + i, \pi + 9i,$$

$$Re(3 + 5i) = 3 \text{ and } Im(3 + 5i) = 5.$$

Definition 1.2.4. *(Equality of complex numbers.)*

Two complex numbers $z = x + yi$ and $z' = x' + y'i$, where x, x', y and y' are real numbers, are said to be equal if

$$x = x' \text{ and } y = y'.$$

That is, $Re(Z) = Re(z')$ and $Im(Z) = Im(z')$.

Example 1.2.5. *Suppose x and y are real numbers such that the two complex number $(2x - 3) + 5i$ and $(x + 7) - (y + 1)i$ are equal. Find the values of x and y.*

Solution

Remark 1.2.6.

(1) We identify every real number x as a complex number $x = x + 0i$. In view of this we may think of the set of real number as a subset of the set of complex numbers, i.e., $\mathbb{R} \subseteq \mathbb{C}$.

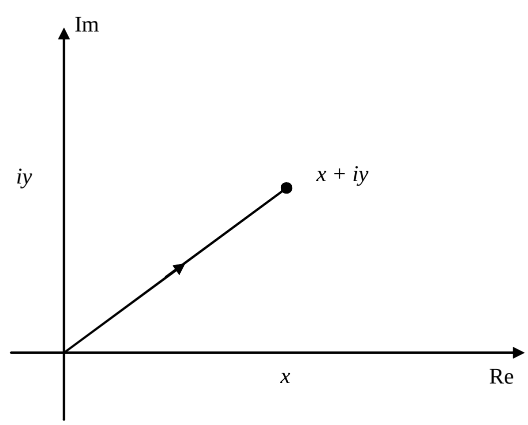
(2) We say that a complex number $z = x + iy$ is purely imaginary if the real part of z, namely x, is zero.

1.3 Argand diagram & Polar representation

Argand Diagram.

1. Since each complex number z is determined uniquely by its real and imaginary parts, we can represent each complex number by a unique point on the xy-place, i.e., by identifying each complex number $z = x + yi$ by the point with coordinate (x, y):

   We can also view each $z = x + iy$ on the Argand diagram as a vector with initial point $(0, 0)$ and terminal point (x, y).



Modulus and Argument

1. From the Argand diagram, the vector (x, y) representing a complex number $z = x + iy$ can also be determined by stating the distance from the point (x, y) to the origin, together with the angle the line segment from $(0, 0)$ to (x, y) made with *the positive x-axis in the anticlockwise direction.*

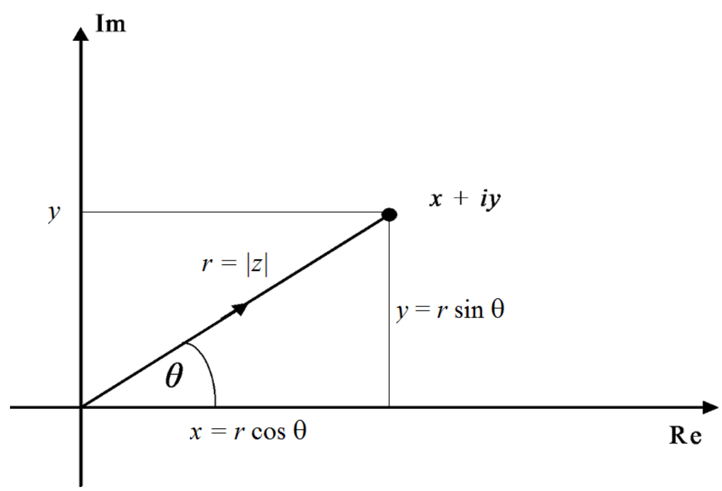
   We shall define such distance the **modulus of** z, denoted by $|z|$, and such an angle an **argument of** z, denoted by $\arg(z)$. So, the modulus $|z|$ of the complex number $z = x + yi$ is $|z| = \sqrt{x^2 + y^2}$ and its argument $\arg(z)$ is the angle θ such that

$$x = |z|\cos\theta \quad \text{and} \quad y = |z|\sin\theta.$$

Note that $\tan\theta = \frac{y}{x}$, if $x \neq 0$.

If $\arg(z) = \theta$,then $\arg(z) = \theta + 2k\pi$, for every integer k.

In particular, when the angle θ is chosen such that $-\pi < \theta \leq \pi$, we say this is the **principal argument** of z.



2. Polar form of z

   Using the modulus and argument we can express a complex number $z = x + iy$ as $z = r(\cos\theta + i\sin\theta)$ or $z = rcis\theta$, where $r = |z|$ and θ is an argument of z. This representation is known as the *polar form* (also known as *trigonometric form*) of z.

3. Exponential form of z

The **exponential form** of a complex number $z = r(\cos\theta + i\sin\theta)$ is $z = re^{i\theta}$. This is commonly used in electronics, engineering and physics. It is convenient in discussing multiplication, division of complex numbers.

Example 1.3.1. *Let $z = 3 - 3i$.*

(a) Find the modulus and principal argument of z, and hence find its polar representation.

(b) Write down the exponential form of z.

Solution

Example 1.3.2. *Express $z = 5e^{\frac{-5\pi}{3}i}$ in rectangular form.*

Solution

1.4 Operations on complex numbers

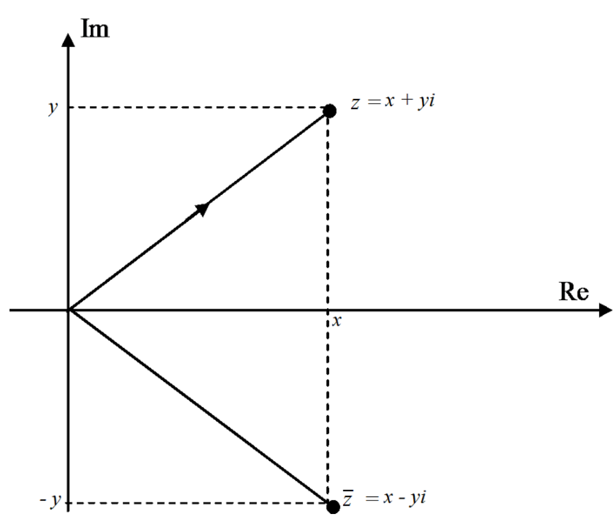
In this section, we discuss taking conjugation of a complex number, and four arithmetic operations between two complex numbers. We shall also look at the representation of these operations in Argand diagram. We also state and discuss related properties.

1.4.1 Conjugate of a complex number.

Definition 1.4.1. *The **conjugate** of a complex number $z = x + yi$ is the complex number $\bar{z} = x - yi$.*

Other notation for the complex conjugate of z is z^*.

Argand Diagram representing z and $\bar{z}$:



Example 1.4.2. *Write down the conjugates of each of the following complex numbers:*

z	$\bar{z}$
$3 + 5i$	
10	
$3.5 - i$	
$-\sqrt{3} + i$	
	$\pi + 9i$
$-\sqrt{7}i$	

Conjugate in Polar Form

The conjugate of the complex number

$z = r(\cos\theta + i\sin\theta)$ (in polar form) or $z = re^{i\theta}$ (in exponential form), is respectively

$$\begin{aligned} z^* &= r(\cos(-\theta) + i\sin(-\theta)), \text{ or} \\ z^* &= re^{-i\theta}. \end{aligned}$$

Proposition 1.4.3. *(a) For every complex number z, we have $\overline{(\overline{z})} = (z^*)^* = z$.*

(b) A complex number z is real if and only of $z = \bar{z}$.

(c) A complex number z is imaginary if and only of $z = -\bar{z}$.

(d) For every complex number z, note that $|\overline{z}| = |z|$ and $\arg(\overline{z})) = -\arg(z)$.

1.4.2 Addition & Subtraction

Complex numbers can be added or subtracted by adding or subtracting the real parts and the imaginary parts separately. Formally we define the two operations as follows:

Definition 1.4.4. *Given two complex numbers $z_1 = x_1 + y_1 i$ and $z_2 = x_2 + y_2 i$, we define*

$$z_1 \pm z_2 = (x_1 + y_1 i) \pm (x_2 + y_2 i) = (x_1 \pm x_2) + (y_1 \pm y_2)i.$$

Example 1.4.5.

(a) $(3 + 5i) + (3.5 - i) = 6.5 - 4i$

(b) $(-\sqrt{3} + i) - (\pi + 9i) = (-\sqrt{3} - \pi) + (-8)i$

Proposition 1.4.6.

(i) $z + 0 = z = 0 + z$, *where* $0 = 0 + 0i$ *is the real number zero.*

(ii) For every $z = x + iy$, *the complex number* $-z = -x + (-y)i$ *satisfies* $z + (-z) = 0 = (-z_1) + z$.

(iii) $z_1 + z_2 = z_2 + z_1$ *(Commutative Law for Addition).*

(iv) $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$ *(Associative Law for Addition).*

(v) $\overline{z_1 \pm z_2} = \overline{z_1} \pm \overline{z_2}$.

1.4.3 Multiplication

To multiply two complex numbers $z_1 = x_1 + y_1 i$ and $z_2 = x_2 + y_2 i$, we can perform the multiplication treating i as a symbol. But we replace i^2 by (-1) when we simplify it :

$$\begin{aligned} z_1 \cdot z_2 &= (x_1 + y_1 i)(x_2 + y_2 i) \\ &= x_1 x_2 + x_1 y_2 i + (y_1 i) x_2 + (y_1 i)(y_2 i) \\ &= (x_1 x_2 - y_1 y_2) + (x_1 y_2 + x_2 y_1) i \end{aligned}$$

Example 1.4.7. *Evaluate* $(3+5i)\cdot(2-i)$

Solution

Proposition 1.4.8. *(i)* $z \cdot 1 = z = 1 \cdot z$.

(ii) $z_1 \cdot z_2 = z_2 \cdot z_1$ *(Commutative Law for product).*

(iii) $(z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3)$ *(Associative Law for product).*

(iv) $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$.

(v) $z \cdot \overline{z} = |z|^2$. *In particular, if* $z \neq 0$, *then* $z \cdot \overline{z} > 0$.

(vi) $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$ *(Distributive Property)*

Product in Polar Form

From the above algebraic expression, it is not clear there is a geometrical relationship between z_1, z_2 and their product. We shall express the product in polar representation to deduce a geometrical relation.

Let us express the two complex numbers in polar form:

$z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$.

We have

$$\begin{aligned} z_1 \cdot z_2 &= r_1(\cos\theta_1 + i\sin\theta_1)\cdot r_2(\cos\theta_2 + i\sin\theta_2) \\ &= \end{aligned}$$

$$= r_1 r_2 \left(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\right).$$

Proposition 1.4.9.

1. *Modulus of the product is the product of moduli:*

$$|z_1 \cdot z_2| = r_1 r_2 = |z_1|\ |z_2|$$

2. *Argument of the product is the sum of arguments:*

$$\arg(z_1 \cdot z_2) = \theta_1 + \theta_2 = \arg(z_1) + \arg(z_2).$$

This implies the complex number $z_1 \cdot z_2$ lies on the line obtained by rotating the line segment representing z_1 by the angle $\arg(z_2)$.

Represent the product of z_1 and z_2 on an Argand diagram:

Remark For a complex number z, the complex number $z \cdot (\cos\theta + i\sin\theta)$ is represented on the Argand diagram by by rotating z through θ.

1.4.4 Division

Recall that to express $\dfrac{1}{3+2\sqrt{5}}$ in the form $a + b\sqrt{5}$, we perform the following

$$\frac{1}{3+2\sqrt{5}} \cdot \frac{3-2\sqrt{5}}{3-2\sqrt{5}} = \frac{3-2\sqrt{5}}{3^2+(2\sqrt{5})^2} = \frac{3}{29} - \frac{2}{29}\sqrt{5}.$$

In a similar way, to divide a complex number $z_1 = x_1 + y_1 i$ by a **non-zero** complex number $z_2 = x_2 + y_2 i$ (i.e., $z_2 \neq 0$), we use the conjugate $\bar{z}_2 = x_2 - y_2 i$, where $z_2 \cdot \bar{z}_2 = x_2^2 + y_2^2$ is a positive real number.

We have

$$\frac{z_1}{z_2} = \frac{z_1}{z_2} \cdot \frac{\overline{z_2}}{\overline{z_2}}$$

$$= \frac{z_1 \overline{z_2}}{z_2 \overline{z_2}} = \frac{z_1 \overline{z_2}}{|z_2|^2}$$

Example 1.4.10. *Evaluate* $\frac{3+5i}{2-i}$.

Solution

Division in Polar Form

In polar form, we have $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$, such that

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$$

$$\arg(\frac{z_1}{z_2}) = \arg(z_1) - \arg(z_2)$$

Thus, we have

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}\left(\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)\right).$$

Example 1.4.11. *Let $z = \cos\theta + i\sin\theta$. Find $|z|$ and show that $\frac{1}{z} = \bar{z}$.*

Solution

1.5 The Fundamental Theorem of Algebra

Theorem 1.5.1 (The Fundamental Theorem of Algebra)**.** *Every polynomial equation of the form*

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0,$$

in which the coefficients $a_0, a_1, \ldots, a_{n-1}, a_n$ are any complex numbers, whose degree n is greater than or equal to one, and whose leading coefficient a_n is not zero, has exactly n roots in the complex number system, provided each multiple root of multiplicity m is counted as m roots.

Proof (Omitted): Textbook on theory of complex analysis.

1.5.1 Solving irreducible quadratic equation

Recall that for a quadratic equation $ax^2 + bx + c = 0$ (or simply a quadratic expression), its **discriminant** ,D, is defined as $D = b^2 - 4ac$.

(i) If $D > 0$, the quadratic equation $ax^2 + bx + c = 0$ has two distinct real roots given by

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

(ii) If $D = 0$, the quadratic equation $ax^2 + bx + c = 0$ has repeat real roots given by

$$x = \frac{-b}{2a}.$$

(iii) If $D < 0$, the quadratic equation $ax^2 + bx + c = 0$ has two distinct complex roots given by

$$x = \frac{-b \pm i\sqrt{-(b^2 - 4ac)}}{2a}.$$

Note that the two complex roots are conjugate of each other.

When $D < 0$, the quadratic equation or expression is said to be irreducible.

Example 1.5.2. *Solve the quadratic equation* $2x^2 - 3x + 5 = 0$.

Solution

In fact there is a more general result, stated below, which we are ready to prove.

Polynomial with Real Coefficients

Theorem 1.5.3. *Suppose $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ is a polynomial in x with real coefficients a_k's. If z is a solution to $p(x) = 0$, then the conjugate $\bar{z}$ is also a solution of $p(x) = 0$.*

For example: suppose we are given that $z_0 = 1+i$ is a complex root of $x^3 - x^2 + 2 = 0$, then ____________ is also a complex root of $x^3 - x^2 + 2 = 0$. Therefore, ______________________________ is a quadratic factor of $x^3 - x^2 + 2$. Moreover, $(x - z_0)(x - \bar{z}_0) = x^2 - (z_0 + \bar{z}_0)x + z\bar{z}_0$ is a real coefficient quadratic factor.

As a consequence of the Fundamental Theorem of Algebra and the above result, we have the following useful result.

Theorem 1.5.4. *Every odd degree polynomial $p(x)$ with real coefficients has at least one real root.*

For example: $9x^5 + 7x^2 - 6x + \pi = 0$ has at least one real root.

1.6 De Moivre's Theorem

In this section, we shall state and prove de Moivre's Theorem. For a positive integer n, we use De Moivre's Theorem to find n-th root of a complex number. In particular, we apply it to determine all distinct n-th roots of unity.

Theorem 1.6.1 (De Moivre's Theorem)**.**

For every integer n,

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta.$$

We look at some examples to illustrate the theorem before we discuss its proof.

Example 1.6.2. *Express each of the following complex numbers in the form $\cos n\theta + i\sin n\theta$.*

(a) $(\cos\theta + i\sin\theta)^9$

(b) $(\cos\theta + i\sin\theta)^{-4}$

Solution

Example 1.6.3. *Simplify each of the following complex numbers:*

(a) $(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4})^{-2}$

(b) $(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3})^{9}$

Example 1.6.4. *Express each of the following complex numbers in the form* $(\cos\theta + i\sin\theta)^n$*.*

(a) $\cos 7\theta + i\sin 7\theta$*.*

(b) $\cos 5\theta - i\sin 5\theta$

Solution

PROOF of De Moivre's Theorem

We prove the theorem by considering two cases: First Case: n is a non-negative integer, i.e., $n \geq 0$ and Second Case: n is a negative integer, i.e., $n < 0$.

Case 1 n is non-negative integer

We shall prove

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta,$$

for $n = 0, 1, 2, 3, \ldots$ by Mathematical Induction.

1. Verify the result holds for $n = 0$

$$(\cos\theta + i\sin\theta)^0 = 1, \cos 0\theta + i\sin 0\theta = \cos 0 = 1.$$

2. Assume the result hold for some non-negative integer k

$$(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta.$$

3. We shall prove the result holds for $k+1$ i.e.,

$$(\cos\theta + i\sin\theta)^{k+1} = \cos(k+1)\theta + i\sin(k+1)\theta.$$

Indeed:

$$\begin{aligned}
\text{LHS} = (\cos\theta + i\sin\theta)^{k+1} &= (\cos\theta + i\sin\theta)^k(\cos\theta + i\sin\theta) \\
&= (\cos k\theta + i\sin k\theta)(\cos\theta + i\sin\theta) \\
&= (\cos k\theta\cos\theta - \sin k\theta\sin\theta) + i(\sin k\theta\cos\theta + \cos k\theta\sin\theta) \\
&= \cos(k+1)\theta + i\sin(k+1)\theta.
\end{aligned}$$

Therefore by Mathematical induction, $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ for all **non-negative integer** n.

Case 2 n is a negative integer, i.e., $n = -1, -2, -3, \ldots$.

Let $n = -m$ where m is a positive integer. Note that

$$(\cos\theta + i\sin\theta)^n = (\cos\theta + i\sin\theta)^{-m}$$

$$= \frac{1}{(\cos\theta + i\sin\theta)^m} = \frac{1}{\cos m\theta + i\sin m\theta}$$

$$= \frac{1}{\cos m\theta + i\sin m\theta}\cdot\frac{\cos m\theta - i\sin m\theta}{\cos m\theta - i\sin m\theta}$$

$$= \frac{\cos m\theta - i\sin m\theta}{\cos^2(m\theta) + \sin^2(m\theta)} = \cos m\theta - i\sin m\theta$$

$$= \cos(-m\theta) + i\sin(-m\theta) = \cos n\theta + i\sin n\theta.$$

1.7 Finding nth roots of $z = r(\cos\alpha + i\sin\alpha)$

We begin with an example to have a geometrical idea of finding roots of a complex number.

Example 1.7.1. *Find all distinct cube roots of* $z = \cos\frac{\pi}{3} + i\sin\frac{\pi}{3}$

The cube roots of $\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}$ are:

$$(\cos\frac{\pi}{9} + i\sin\frac{\pi}{9}),\ (\cos\frac{7\pi}{9} + i\sin\frac{7\pi}{9})\ \ \&\ \ (\cos\frac{13\pi}{9} + i\sin\frac{13\pi}{9}).$$

Theorem 1.7.2 (Distinct nth roots)**.**

Consider a complex number z in polar form

$$z = r(\cos\alpha + i\sin\alpha), \textit{ where } r > 0 \textit{ and } -\pi < \alpha \leq \pi.$$

The nth distinct roots of the complex number $z = r(\cos\alpha + i\sin\alpha)$ are

$$z_k = \sqrt[n]{r}\left(\cos\frac{2k\pi + \alpha}{n} + i\sin\frac{2k\pi + \alpha}{n}\right), k = 0, 1, 2, 3, \ldots, n-1.$$

Theorem 1.7.3 (Distinct nth roots - exponential form)**.**

In exponential form, we have all n distinct nth roots of the complex number $z = re^{i\alpha}$ are

$$z_k = \sqrt[n]{r}\left(e^{i\frac{\theta + 2k\pi}{n}}\right), k = 0, 1, 2, 3, \ldots, n-1.$$

The n integers can be chosen to be any n consecutive integers.

Example 1.7.4. ***Find all*** 5***th roots of*** $\sqrt{3} + i$

Solution

Corollary 1.7.5. *The n distinct complex numbers*

$$w_k = cis\ (\frac{\alpha}{n} + \frac{2k\pi}{n}) = \cos(\frac{\alpha}{n} + \frac{2k\pi}{n}) + i\sin(\frac{\alpha}{n} + \frac{2k\pi}{n}), k = 0, 1, 2, \ldots, n-1$$

are all n distinct roots of $\cos\theta + i\sin\theta$.

Example 1.7.6. *The nine distinct 9^{th} roots of $\cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i$ are*

$$w_k = cis\ (\frac{\pi/2}{9} + \frac{2k\pi}{9}) = cis\ (\frac{\pi + 4k\pi}{18}), k = 0, 1, 2, 3, \ldots, 8,$$

They are

$$cis\ (\frac{\pi}{18}), cis\ (\frac{5\pi}{18}), cis\ (\frac{9\pi}{18}) = cis\ \frac{\pi}{2} = i, cis\ (\frac{13}{18}), cis\ (\frac{17\pi}{18}), cis\ (\frac{21\pi}{18}), cis\ (\frac{25\pi}{18}), cis\ (\frac{29}{18})\ \ \&\ \ cis\ (\frac{33}{18})$$

1.7.1 Finding nth roots of unity.

Note that $1 = 1 + 0i = \cos 0 + i \sin 0 = \cos 2k\pi + i \sin 2k\pi$, where k is an integer. We call n-th roots of 1 the n-th roots of unity.

Corollary 1.7.7 (nth roots of unity)**.** *The n distinct nth roots of unity are*

$$z_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, k = 0, 1, 2, 3, \ldots, n-1.$$

By De Moivre's Theorem, we have $z_k = (z_1)^k$.

On the Argand diagram, all n-th roots of 1 are represented by points on the unit circle and they are equally spaced by $\frac{2\pi}{n}$:

1.7.2 Deriving Certain Trigonometric Identities I

We can use De Moivre's Theorem to express $\cos n\theta$, $\sin n\theta$ and $\tan n\theta$ in terms of powers of $\cos\theta$, $\sin\theta$ and $\tan\theta$.

Tools:

$$\cos n\theta = \mathrm{Re}(\cos n\theta + i\sin n\theta) = \mathrm{Re}\left(\cos\theta + i\sin\theta\right)^n,$$
$$\sin n\theta = \mathrm{Im}(\cos n\theta + i\sin n\theta) = \mathrm{Im}\left(\cos\theta + i\sin\theta\right)^n,$$

Apply binomial expansion to $(\cos\theta + i\sin\theta)^n$

Notation used: $c \equiv \cos\theta$, $s \equiv \sin\theta$, $t \equiv \tan\theta$.

Example 1.7.8. *Express* $\sin 3\theta$ *in terms of powers of* $\sin\theta$.

The first step is to note that

$$\sin 3\theta = \mathrm{Im}(\cos 3\theta + i\sin 3\theta)$$

Now, we apply de Moivre's theorem

$$\begin{aligned}
\sin 3\theta &= \mathrm{Im}(\cos 3\theta + i\sin 3\theta) \\
&= \mathrm{Im}\left(\cos\theta + i\sin\theta\right)^3 \text{ (why?)} \\
&= \mathrm{Im}\left(c + is\right)^3 \\
&= \mathrm{Im}\left(c^3 + 3c^2 is + 3ci^2s^2 + i^3s^3\right) \\
&= \mathrm{Im}\left(c^3 - 3cs^2 + i(3c^2s - s^3)\right) \\
&= 3c^2s - s^3
\end{aligned}$$

Using $c^2 + s^2 = 1$, we have

$$\begin{aligned}
\sin 3\theta &= 3c^2 s - s^3 \\
&= 3(1-s^2)s - s^3 \\
&= 3s - 4s^3 \\
&= 3\sin\theta - 4\sin^3\theta.
\end{aligned}$$

From the above, we have also obtained an expression for $\cos 3\theta$:

$$\cos 3\theta = c^3 - 3cs^2 = c^3 - 3c(1-c^2) = 4c^3 - 3c$$

Using the expression for both $\sin 3\theta$ and $\cos 3\theta$, we obtain a similar expression for $\tan 3\theta$:

$$\tan 3\theta = \frac{\sin 3\theta}{\cos 3\theta} = \frac{3c^2 s - s^3}{c^3 - 3cs^2}$$
$$= \frac{3c^2 s - s^3}{c^3 - 3cs^2}.\left(\frac{1/c^3}{1/c^3}\right) = \frac{3t - t^3}{1 - 3t^2}$$

1.7.3 Deriving Certain Trigonometric Identities II

Aim: to express $\cos^n\theta$ or $\sin^n\theta$ in terms of cosines and sines of multiples of θ, i.e. $\cos k\theta$, $\sin k\theta$.

Main Tool: Let $z = \cos\theta + i\sin\theta$, we have $\frac{1}{z} = \cos\theta - i\sin\theta$.

Thus we have $\cos\theta = \frac{1}{2}\left(z + \frac{1}{z}\right)$ and $\sin\theta = \frac{1}{2i}\left(z - \frac{1}{z}\right)$. Next, we apply binomial expansion and group z^k and $\frac{1}{z^k}$ together.

By De Moivre's Theorem, we have

$$z^k = \cos k\theta + i\sin k\theta \text{ and } \frac{1}{z^k} = \cos k\theta - i\sin k\theta$$

which gives

$$z^k + \frac{1}{z^k} = 2\cos k\theta \text{ and } z^k - \frac{1}{z^k} = 2i\sin k\theta.$$

Thus, we obtain an expression involving sines and cosines of multiple of θ.

Example 1.7.9. *Prove that* $\cos^3\theta = \frac{1}{4}(\cos 3\theta + 3\cos\theta)$

Proof. Let $z = \cos\theta + i\sin\theta$. We have

$$\cos^3\theta = (\cos\theta)^3 = \left(\frac{1}{2}(z + \frac{1}{z})\right)^3$$

$$
\begin{aligned}
&= \frac{1}{8}\left(z^3 + 3z^2\frac{1}{z} + 3z(\frac{1}{z})^2 + (\frac{1}{z})^3\right) = \frac{1}{8}\left((z^3 + \frac{1}{z^3}) + 3(z + \frac{1}{z})\right) \\
&= \frac{1}{8}\left[2\cos 3\theta + 3(2\cos\theta)\right] = \frac{1}{4}\left(\cos 3\theta + 3\cos\theta\right).
\end{aligned}
$$

□

Chapter 2

Vectors

Vectors are quantities which have both magnitude and direction. Examples of vectors include acceleration, displacement, force, momentum and velocity.

2.1 Geometrical Representation of Vectors

We may use a **directed** segment to represent a vector graphically. The directed segment from a point A to a point B is denoted by $\overrightarrow{AB}$. The length $|AB|$ of the segment represents the magnitude of the vector.

We may also use lower case bold letters **u, v, w** to denote vectors.

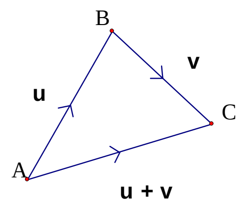
- Two vectors are equal if they have the same magnitude and direction.
- Given a vector $\mathbf{v}$, its negative $-\mathbf{v}$ is the vector with the same magnitude and opposite direction.
- The **zero vector 0** is a vector with zero magnitude.

2.2 Vector Addition & Scalar Multiplication

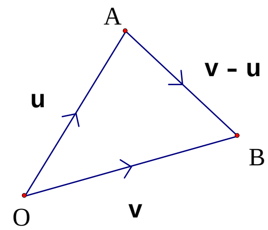
Vector Addition and Subtraction

Given two vectors $\mathbf{u}$ and $\mathbf{v}$, we add them to form a new vector $\mathbf{u} + \mathbf{v}$ as follows:

If $\mathbf{u} = \overrightarrow{AB}$ and $\mathbf{v} = \overrightarrow{BC}$, then $\mathbf{u} + \mathbf{v} = \overrightarrow{AC}$. Or simply $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.



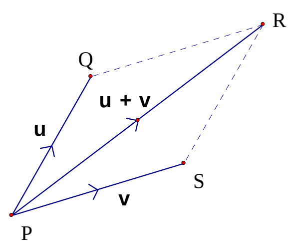
Vector Addition



Vector Subtraction

If $\mathbf{u} = \overrightarrow{OA}$ and $\mathbf{v} = \overrightarrow{OB}$, then $\overrightarrow{AB} = \mathbf{v} - \mathbf{u}$. Or simply $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$.

Parallelogram Law of Vector Addition



Scalar multiple of vector

If $\lambda > 0$, then $\lambda\mathbf{v}$ is a vector in the direction of $\mathbf{v}$ with magnitude λ times that of $\mathbf{v}$.

Conversely, any vector that has the same direction as $\mathbf{v}$ must be equal to $\lambda\mathbf{v}$ for some $\lambda > 0$.

Properties of vector addition and scalar multiplication

- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$,

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

- $\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$, $\lambda, \mu \in \mathbb{R}$,

- $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$, $\lambda, \mu \in \mathbb{R}$,

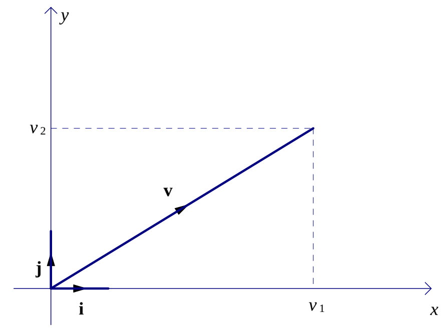
- $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \boldsymbol{\lambda}\mathbf{v}$, $\lambda \in \mathbb{R}$.

2.3 Vectors in Coordinate System

2.3.1 Vectors in 2-space (the plane).

We may position a vector $\mathbf{v}$ in the plane $\mathbb{R}^2$ with its initial point at the origin O and the terminal point (v_1, v_2) and we write $\mathbf{v} = (v_1, v_2)$, or $\mathbf{v} = u_1\mathbf{i} + v_2\mathbf{j}$, where $\mathbf{i}$ and $\mathbf{j}$ are unit vectors in the positive direction of x and y axis respectively.

The vector $\mathbf{v}$ may be written in the column vector form $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

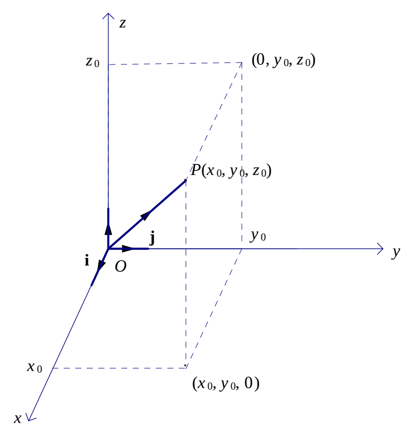


Suppose $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$

1. $\mathbf{u} = (u_1, u_2)$ is the zero vector if and only if $u_1 = u_2 = 0$.
2. Two vectors $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ are equal if and only if $u_1 = v_1$ and $u_2 = v_2$.
3. $\mathbf{u} + \mathbf{v} = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$.
4. $\lambda\mathbf{u} = \lambda (u_1, u_2) = (\lambda u_1, \lambda u_2), \lambda \in \mathbb{R}$.

2.3.2 Vectors in 3-space

Consider a rectangular coordinate system, with three mutually perpendicular lines, called coordinate axes, passing through a point called the origin.



Note that we use a right-handed system:

With your palm of right hand along the positive x-axis, rotate $90°$ towards the positive y-axis, your thumb points in the positive z-axis.

Like vectors in 2-space, a vector in 3-space from the point $P_1 = (x_1, y_1, z_1)$ to the point $P_2 = (x_2, y_2, z_2)$ will be the vector $\overrightarrow{P_1P_2} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$. In particular, if the starting point is the origin $O = (0, 0, 0)$ and the end point $P = (x_0, y_0, z_0)$, then $\overrightarrow{OP} = (x_0 - 0, y_0 - 0, z_0 - 0) = (x_0, y_0, z_0)$.

The numbers x_0, y_0 and z_0 are called the components of $\mathbf{v}$.

We also write

$$\mathbf{v} = x_0\mathbf{i} + y_0\mathbf{j} + z_0\mathbf{k},$$

where $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are unit vectors in the positive direction of x, y and z axis respectively. The vector $\mathbf{v}$ may be written in the column vector form $\mathbf{v} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$.

Suppose $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (u_1, v_2, v_3)$ are two vectors in 3-space. Then we have the following:

1. $\mathbf{u} = (u_1, u_2, u_3)$ is the zero vector if and only if $u_1 = 0$, $u_2 = 0$ and $u_3 = 0$.

2. Two vectors $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (u_1, v_2, v_3)$ are equal if and only if $u_1 = v_1$, $u_2 = v_2$ and $u_3 = v_3$.

3. $\mathbf{u} + \mathbf{v} = (u_1, u_2, u_3) + (u_1, v_2, v_3) = (u_1 + u_1, u_2 + v_2, u_3 + v_3)$.

4. $k\mathbf{u} = k(u_1, u_2, u_3) = (ku_1, ku_2, ku_3)$.

Length or norm of a vector

In 2-space, the length, (or norm, magnitude) of a vector $\mathbf{u} = (u_1, u_2)$ is $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2}$.

In 3-space, the length of a vector $\mathbf{u} = (u_1, u_2, u_3)$ is $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2}$.

Example 2.3.1. *Evaluate the norm of following vectors:*

(a) $\mathbf{u} = (2, -5)$ *in* 2*-space.*

(b) $\mathbf{v} = (-1, 2, 2)$ *in* 3*-space.*

(c) $\mathbf{u} = (\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$ *in* 3*-space.*

Solution

Remark The distance d between two points $P = (u_1, u_2, u_3)$ and $Q = (u_1, v_2, v_3)$ is given by the norm $\|\overrightarrow{PQ}\|$ of the vector $\overrightarrow{PQ}$,

$$d = \sqrt{(v_1 - u_1)^2 + (v_2 - u_2)^2 + (v_3 - u_3)^2}.$$

Note that it follows from the definition of $k\mathbf{u}$ that $\|k\mathbf{u}\| = |k|\|\mathbf{u}\|$.

Unit Vectors

A vector of length 1 is called a **unit vector**. The vector $\mathbf{u} = (-1/\sqrt{2}, 0, 1/\sqrt{2})$ is a unit vector.

In particular, the following vectors

$$\mathbf{i} = (1, 0, 0), \mathbf{j} = (0, 1, 0) \text{ and } \mathbf{k} = (0, 0, 1)$$

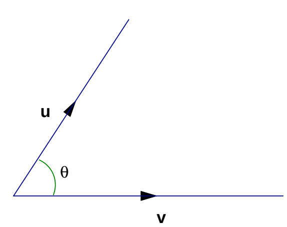
are unit vectors along the positive direction of x-, y- and z-axes respectively.

It follows that:

For a nonzero vector $\mathbf{u}$, the vector $\hat{\mathbf{u}} = \frac{1}{\|\mathbf{u}\|}\mathbf{u}$ is the unit vector along the direction $\mathbf{u}$.

Dot product and projections

The *dot product* (or scalar product) of two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$ is defined by $\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}||\,||\mathbf{v}||\cos\theta$,
where $\theta \in [0, \pi]$ is the angle between the two vectors.
If either of $\mathbf{u}$ or $\mathbf{v}$ is a zero vector, then define $\mathbf{u} \cdot \mathbf{v} = 0$.



Example 2.3.2. *Find the dot product of each pair of vectors.*

(a) $\mathbf{u} = (1, -1)$ and $\mathbf{v} = (2, 0)$.

(b) $\mathbf{u} = (1, -1)$ and $\mathbf{w} = (0, 2)$.

(c) $\mathbf{u} = (1, -1)$ and $\mathbf{s} = (2, 2)$.

Solution

Properties of dot product

1. The dot product of two vectors is a real number (scalar).
2. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}.$,
3. For any vector $\mathbf{u}$, $\mathbf{u} \cdot \mathbf{u} = ||\mathbf{u}||^2$,
4. Suppose that $\mathbf{u}$ and $\mathbf{v}$ are non-zero vectors, then $\mathbf{u} \cdot \mathbf{v} = 0 \Leftrightarrow \mathbf{u} \perp \mathbf{v}$.
5. $(\lambda \mathbf{u}) \cdot (\mu \mathbf{v}) = (\lambda \mu)\, \mathbf{u} \cdot \mathbf{v}$.
6. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{c}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{c}$.
7. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{c} = \mathbf{u} \cdot \mathbf{c} + \mathbf{v} \cdot \mathbf{c}$.

Dot product of vectors in column vectors form

Proposition 2.3.3. *If* $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ *and* $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, *then* $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2$.

Proof. Since $\mathbf{u} = u_1 \mathbf{i} + u_2 \mathbf{j}$ and $\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j}$,

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= (u_1 \mathbf{i} + u_2 \mathbf{j}) \cdot (v_1 \mathbf{i} + v_2 \mathbf{j}) \\ &= (u_1 v_1)\, \mathbf{i} \cdot \mathbf{i} + (u_1 v_2)\, \mathbf{i} \cdot \mathbf{j} + (u_2 v_1)\, \mathbf{j} \cdot \mathbf{i} + (u_2 v_2)\, \mathbf{j} \cdot \mathbf{j} \\ &= u_1 v_1 \, ||\mathbf{i}||^2 + u_2 v_2 \, ||\mathbf{j}||^2, \text{ since } \mathbf{i} \cdot \mathbf{j} = 0 \; (\mathbf{i} \perp \mathbf{j}) \\ &= u_1 v_1 + u_2 v_2. \end{aligned}$$

□

Similarly, we can show that:

If $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$, then $\underline{\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3.}$

Example 2.3.4. *Find the* $\mathbf{u} \cdot \mathbf{v}$ *if* $\mathbf{u} = (2, 3, -1)$ *and* $\mathbf{v} = (1, 0, 5)$.

Solution We have

$$\mathbf{u}\cdot\mathbf{v} = (2,3,-1)\cdot(1,0,5) = 2(1)+3(0)+(-1)(5) = -3.$$

Since $\mathbf{u}\cdot\mathbf{v} < 0$, the angle θ between the two vectors is an obtuse angle.

Angle between two vectors

The above formula for dot product turns out to be very useful, especially in finding the angle between two given vectors.

If θ is the angle between two vectors $\mathbf{u}$ and $\mathbf{v}$, then

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{||\mathbf{u}||\,||\mathbf{v}||}.$$

Example 2.3.5. *Find the angle between* $\mathbf{u} = \mathbf{i}+2\mathbf{j}-2\mathbf{k}$ *and* $\mathbf{v} = 2\mathbf{i}+3\mathbf{j}-6\mathbf{k}$.

Solution Let θ denote the angle between $\mathbf{u}$ and $\mathbf{v}$.

$\mathbf{u}\cdot\mathbf{v} = (1)\,(2)+(2)\,(3)+(-2)\,(-6) = 20.$

$$||\mathbf{u}|| = \sqrt{\left(\quad\right)^2 + \left(\quad\right)^2 + \left(\quad\right)^2} = 3$$

$$||\mathbf{v}|| = \sqrt{\left(\quad\right)^2 + \left(\quad\right)^2 + \left(\quad\right)^2} = 7$$

Then

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{||\mathbf{u}||\,||\mathbf{v}||}$$

$$=$$

$$= \frac{20}{21}.$$

$$\theta = \cos^{-1}\frac{20}{21} = 0.310 \text{ rad}$$

Example 2.3.6. *Find the angle between* $\mathbf{u} = -\mathbf{i}-2\mathbf{j}+2\mathbf{k}$ *and* $\mathbf{v} = 2\mathbf{i}+3\mathbf{j}-6\mathbf{k}$.

Solution Let θ denote the angle between $\mathbf{u}$ and $\mathbf{v}$.

$\mathbf{u}\cdot\mathbf{v} = (-1)\,(2)+(-2)\,(3)+(2)\,(-6) = -20.$

$$||\mathbf{u}|| = \sqrt{\left(\quad\right)^2 + \left(\quad\right)^2 + \left(\quad\right)^2} = 3$$

$$||\mathbf{v}|| = \sqrt{\left(\quad\right)^2 + \left(\quad\right)^2 + \left(\quad\right)^2} = 7$$

Then

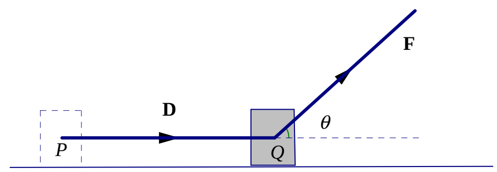
$$\begin{aligned}
\cos\theta &= \frac{\mathbf{u}\cdot\mathbf{v}}{||\mathbf{u}||\,||\mathbf{v}||} \\
&= \\
&= -\frac{20}{21} \ (\cos\theta < 0 \text{ implies that } \theta \text{ is} \qquad\qquad\qquad) \\
\theta &= \pi - \cos^{-1}\frac{20}{21} = \qquad\qquad\qquad \text{rad}
\end{aligned}$$

Work Done

One of the many applications of dot product in science and engineering is to compute work done. Recall that force and displacement are vector quantities, they have magnitude and direction.

The work done by a constant force $\mathbf{F}$ acting through a displacement $\mathbf{D}$ is the dot product

$$W = \mathbf{F} \cdot \mathbf{D}.$$



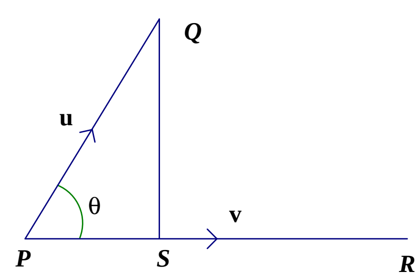
Example 2.3.7. *If* $\|\mathbf{F}\| = 50N$*,* $\|\mathbf{D}\| = 30\ m$*, and* $\theta = 45°$*, then the work done by* $\mathbf{F}$ *acting from* P *to* Q *is given by*

$$W = \mathbf{F} \cdot \mathbf{D} = \|\mathbf{F}\|\ \|\mathbf{D}\| \cos\theta$$

$$= (50)\,(30)\,\frac{1}{\sqrt{2}} = \frac{1500}{\sqrt{2}}\ \mathrm{J}.$$

Projection of a vector

If $\mathbf{u} = \overrightarrow{PQ}$ and $\mathbf{v} = \overrightarrow{PR}$, then $\overrightarrow{PS}$ is the **(perpendicular) projection** of $\mathbf{u}$ onto $\mathbf{v}$.



The length of projection of $\mathbf{u}$ onto $\mathbf{v}$ =

$$\begin{aligned} \left|\overrightarrow{PS}\right| &= \left|\left\|\overrightarrow{PQ}\right\| \cos\theta\right| \\ &= \left| \|\mathbf{u}\| \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} \right| = \left| \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|} \right|. \end{aligned}$$

Then the vector projection of $\mathbf{u}$ onto $\mathbf{v}$, $\mathrm{proj}_{\mathbf{v}}\,\mathbf{u}$

$$= \text{length of projection} \times \hat{\mathbf{v}} = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|}\right)\hat{\mathbf{v}} = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|}\right)\frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

Example 2.3.8. *Find the vector projection of* $\mathbf{u} = \mathbf{i} - 2\mathbf{j} + 2\mathbf{k}$ *onto* $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j} - 6\mathbf{k}$.

Solution

$\mathbf{u} \cdot \mathbf{v} = -16$

$$||\mathbf{u}|| = \sqrt{\left(\qquad\right)^2 + \left(\qquad\right)^2 + \left(\qquad\right)^2} = 3$$

$$||\mathbf{v}|| = \sqrt{\left(\qquad\right)^2 + \left(\qquad\right)^2 + \left(\qquad\right)^2} = 7$$

Therefore $\hat{\mathbf{v}} = \frac{\mathbf{v}}{||\mathbf{v}||} = \frac{2\mathbf{i}+3\mathbf{j}-6\mathbf{k}}{7}$.

Then the vector projection of $\mathbf{u}$ onto $\mathbf{v}$ =

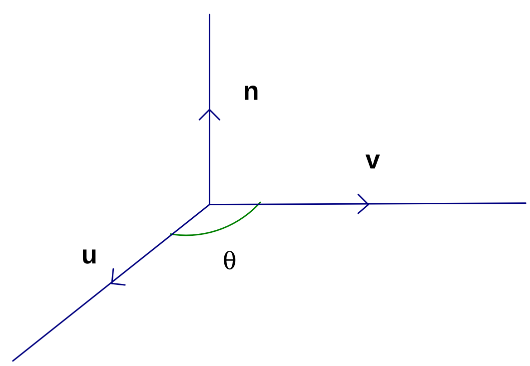
$$\text{proj}_{\mathbf{v}}\, \mathbf{u} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{v}||}\right) \hat{\mathbf{v}}$$

$$=$$

$$= -\frac{16}{49}\left(2\mathbf{i}+3\mathbf{j}-6\mathbf{k}\right).$$

Cross product

The **cross product** (or vector product) of two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined as $\mathbf{u} \times \mathbf{v} = (||\mathbf{u}||\,||\mathbf{v}|| \sin\theta)\, \hat{\mathbf{n}}$, where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$, and $\hat{\mathbf{n}}$ is the vector perpendicular to both $\mathbf{u}$ and $\mathbf{v}$, governed by the right hand rule ($\mathbf{u}$-index finger, $\mathbf{v}$-middle finger, $\mathbf{n}$-thumb, of right hand).



Properties of cross product:

1. The cross product of two vectors is a vector.
2. If $\mathbf{u}$ and $\mathbf{v}$ are parallel, then $\theta = 0$ and thus $\mathbf{u} \times \mathbf{v} = \mathbf{0}$.
3. $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$ (anti-commutative).

4. $\mathbf{u}\times(\mathbf{v}+\mathbf{w})=\mathbf{u}\times\mathbf{v}+\mathbf{u}\times\mathbf{w}$ (distributive wrt addition).

5. $(\lambda\mathbf{u})\times(\mu\mathbf{v})=(\lambda\mu)(\mathbf{u}\times\mathbf{v})$.

Example 2.3.9. $\mathbf{i}\times\mathbf{i}=\mathbf{j}\times\mathbf{j}=\mathbf{k}\times\mathbf{k}=0$

Example 2.3.10. $\mathbf{i}\times\mathbf{j}=\mathbf{k}$; $\mathbf{j}\times\mathbf{k}=\mathbf{i}$; $\mathbf{k}\times\mathbf{i}=\mathbf{j}$

Discriminant formula for cross product

Theorem 2.3.11. *If* $\mathbf{u}=\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}$ *and* $\mathbf{v}=\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}$, *then*

$$\begin{aligned}\mathbf{u}\times\mathbf{v}&=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\u_1&u_2&u_3\\v_1&v_2&v_3\end{vmatrix}\\&=\begin{vmatrix}u_2&u_3\\v_2&v_3\end{vmatrix}\mathbf{i}-\begin{vmatrix}u_1&u_3\\v_1&v_3\end{vmatrix}\mathbf{j}+\begin{vmatrix}u_1&u_2\\v_1&v_2\end{vmatrix}\mathbf{k}.\end{aligned}$$

Proof.

$$\begin{aligned}\mathbf{u}\times\mathbf{v}&=(u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k})\times(v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k})\\&=u_1v_1\mathbf{i}\times\mathbf{i}+u_1v_2\mathbf{i}\times\mathbf{j}+u_1v_3\mathbf{i}\times\mathbf{k}\\&+u_2v_1\mathbf{j}\times\mathbf{i}+u_2v_2\mathbf{j}\times\mathbf{j}+u_2v_3\mathbf{j}\times\mathbf{k}\\&+u_3v_1\mathbf{k}\times\mathbf{i}+u_3v_2\mathbf{k}\times\mathbf{j}+u_3v_3\mathbf{k}\times\mathbf{k}\\&=\mathbf{0}+u_1v_2\mathbf{k}+u_1v_3(-\mathbf{j})\\&+u_2v_1(-\mathbf{k})+\mathbf{0}+u_2v_3(\mathbf{i})\\&+u_3v_1(\mathbf{j})+u_3v_2(-\mathbf{i})+\mathbf{0}\\&=(u_2v_3-u_3v_2)\mathbf{i}-(u_1v_3-u_3v_1)\mathbf{j}+(u_1v_2-u_2v_1)\mathbf{k}\\&=\begin{vmatrix}u_2&u_3\\v_2&v_3\end{vmatrix}\mathbf{i}-\begin{vmatrix}u_1&u_3\\v_1&v_3\end{vmatrix}\mathbf{j}+\begin{vmatrix}u_1&u_2\\v_1&v_2\end{vmatrix}\mathbf{k}.\end{aligned}$$

$\square$

Example 2.3.12. *Find* $\mathbf{u}\times\mathbf{v}$ *if* $\mathbf{u}=(1,1,1)$ *and* $\mathbf{v}=(1,0,1)$.

Solution

$$\begin{aligned}\mathbf{u}\times\mathbf{v}&=(1,1,1)\times(1,0,1)\\&=\left(\begin{vmatrix}1&1\\0&1\end{vmatrix},-\begin{vmatrix}1&1\\1&1\end{vmatrix},\begin{vmatrix}1&1\\1&0\end{vmatrix}\right)\end{aligned}$$

$$= \qquad .$$

Applications of Vector product

1. Perpendicular Vector

The cross product $\mathbf{u} \times \mathbf{v}$ provides a vector perpendicular to both $\mathbf{u}$ and $\mathbf{v}$. From this, we may obtain a unit vector in this direction.

Example 2.3.13. *Find a unit vector* $\hat{\mathbf{n}}$ *which is perpendicular to both* $\mathbf{u} = (1, 1, 1)$ *and* $\mathbf{v} = (1, 0, 1)$.

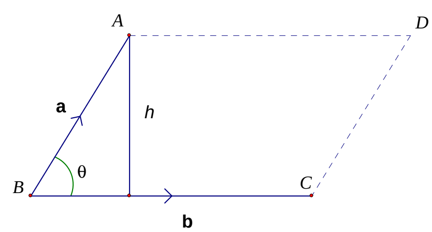
Solution From the previous example,

$$\mathbf{u} \times \mathbf{v} = (1, 0, -1).$$

Thus, we have

$$\hat{\mathbf{n}} = \frac{1}{\sqrt{2}}(1, 0, -1).$$

2. Area of parallelogram $ABCD$
$=$base$\times$height
$= ||BC|| \times h$
$= ||BC|| \, ||AB|| \sin\theta$
$= \left|\left|\overrightarrow{AB} \times \overrightarrow{BC}\right|\right|.$



3. Area of triangle $ABC = \frac{1}{2}$ of Area of parallelogram $ABCD = \frac{1}{2}\left|\left|\overrightarrow{AB} \times \overrightarrow{BC}\right|\right|.$

Example 2.3.14. *Find the area of triangle whose vertices are* $A\,(1, 0, 0)\,,\ B\,(0, 1, 0)$ *and* $C\,(0, 0, 1)$.

Solution Area of triangle $ABC = \frac{1}{2}\left|\left|\overrightarrow{AB} \times \overrightarrow{AC}\right|\right|.$

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \quad \\ \quad \\ \quad \end{pmatrix}$$

$$\overrightarrow{AC} = \overrightarrow{OC} - \overrightarrow{OA} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \quad \\ \quad \\ \quad \end{pmatrix}$$

$$\overrightarrow{AB} \times \overrightarrow{AC} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \quad \\ \quad \\ \quad \end{pmatrix}$$

$$\text{Area of triangle } ABC = \tfrac{1}{2}\left\|\begin{pmatrix} \\ \\ \end{pmatrix}\right\| = \frac{\sqrt{\left(\quad\right)^2 + \left(\quad\right)^2 + \left(\quad\right)^2}}{2} = \frac{\sqrt{3}}{2}.$$

Scalar Triple Product

For three vectors

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \text{ and } \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \text{ in } \mathbb{R}^3,$$

the dot product $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ is called the *scalar triple product* of vectors $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$.

The scalar triple product $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ can be computed numerically as follows:

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = u_1 \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} - u_2 \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} + u_3 \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix}$$

It can be verified that

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})$$

Theorem 2.3.15. *(a) For three non-coplanar $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$, the absolute value $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$ of the scalar triple product is the volume of the parallelepiped which is a three-dimensional figure formed (by six parallelograms) whose sides are the three given vectors.*

(b) For general three vectors $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$, if the scalar triple product $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 0$, we may conclude that either at least one of the vectors $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ is a zero vector or the three vectors are coplanar (they lie on the same plane).

Example 2.3.16. *Find the volume of the parallelepiped formed by vectors $\mathbf{u} = (1, 2, 3)$, $\mathbf{v} = (-1, 1, 0)$ and $\mathbf{w} = (1, 2, 1)$.*

Solution We compute the scalar triple product:

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 1 \begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} - 2 \begin{vmatrix} -1 & 0 \\ 1 & 1 \end{vmatrix} + 3 \begin{vmatrix} -1 & 1 \\ 1 & 2 \end{vmatrix}$$

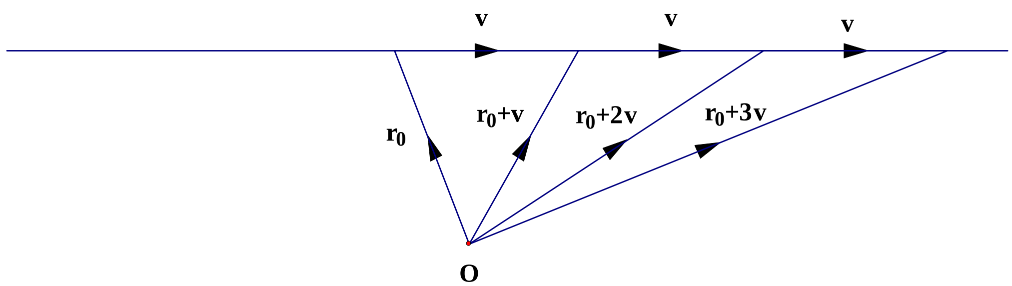
$$= 1 - 2(-1) + 3(-3) = -6$$

Thus the volume of the parallelepiped is 6 unit^3.

2.4 Lines and Planes

2.4.1 Lines

A line on a Cartesian plane may be determined by a point on the line and the gradient of the line. In a space, line is uniquely determined by its **direction vector** and a point on the line.



If a line ℓ is parallel to a vector $\mathbf{v}$ and passes through a point with position vector $\mathbf{r}_0$, then the **vector equation** of ℓ is

$$\ell : \mathbf{r}(t) = \mathbf{r}_0 + t\mathbf{v},\ t \in \mathbb{R}.$$

Example 2.4.1. *If A, B and C have coordinates $(1,1,0)$, $(0,1,0)$ and $(0,0,1)$ respectively. Find a vector equation for the line that is parallel to BC and passes through A.*

Solution

$$\mathbf{r}_0 = \overrightarrow{OA} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\mathbf{v} = \overrightarrow{BC} = \begin{pmatrix} \\ \\ \\ \end{pmatrix} - \begin{pmatrix} \\ \\ \\ \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Therefore, the vector equation of ℓ is

$$\begin{aligned} \mathbf{r}(t) &= \mathbf{r}_0 + t\mathbf{v} \\ &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} \\ \\ \\ \end{pmatrix} \end{aligned}$$

Equation of a line: Cartesian form

Suppose that a line ℓ passes through $\mathbf{r}_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ and is parallel to $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$. If (x, y, z) is a point on ℓ, then

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \; t \in \mathbb{R}.$$

Therefore,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 + tv_1 \\ y_0 + tv_2 \\ z_0 + tv_3 \end{pmatrix}.$$

Thus

$$\underline{x = x_0 + tv_1, \; y = y_0 + tv_2, \; z = z_0 + tv_3.} \tag{2.1}$$

Equation (2.1) is called the *parametric equation* of ℓ.

Suppose that $v_1, v_2, v_3 \neq 0$, we obtain the *Cartesian Equation* of ℓ :

$$\frac{x - x_0}{v_1} = \frac{y - y_0}{v_2} = \frac{z - z_0}{v_3}.$$

Question: What happens if $v_1 = 0$?

Example 2.4.2. If a line ℓ has equation

$$x - 1 = \frac{y + 1}{2} = \frac{z}{3}.$$

(i) Find a vector equation for ℓ. (ii) Find the parametric equation for ℓ.

Solution The Cartesian equation for ℓ :

$$\frac{x - 1}{1} = \frac{y - (-1)}{2} = \frac{z - 0}{3}.$$

(i) The line ℓ is parallel to $(1, 2, 3)$ and passes through $(1, -1, 0)$. Thus the a vector equation for ℓ is

$$\mathbf{r} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \; t \in \mathbb{R}.$$

(ii) The parametric equation for ℓ :

$$x = 1 + t, y = -1 + 2t, z = 3t, t \in \mathbb{R}.$$

Angles between two lines

If ℓ_1 and ℓ_2 have direction vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ respectively and θ is the acute angle between $\mathbf{v}_1$ and $\mathbf{v}_2$, then

$$\theta = \cos^{-1}\left|\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{||\mathbf{v}_1||\,||\mathbf{v}_2||}\right|.$$

Example 2.4.3. *Find the acute angle between z-axis and the line $\ell : x = 1-t, y = 3+\sqrt{2}t$, $z = -5+t$*

Solution. ℓ and z-axis has direction vectors

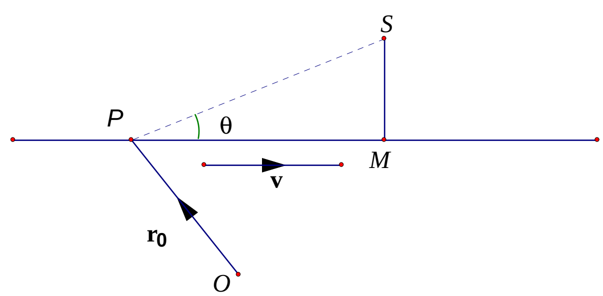
$$\mathbf{v}_1 = \begin{pmatrix} \\ \\ \\ \\ \end{pmatrix} \text{ and } \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

respectively. Let θ be the acute angle between $\mathbf{v}_1$ and $\mathbf{v}_2$. Then

$$\theta = \cos^{-1}\left|\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{||\mathbf{v}_1||\,||\mathbf{v}_2||}\right| = \cos^{-1}\frac{1}{(2)\,(1)} = \frac{\pi}{3} \text{ or } 60^\circ.$$

Distance from a point to a line

If ℓ is a passing through P line parallel to $\mathbf{v}$, S is point not on ℓ and M is the foot of perpendicular of S to ℓ. Then the distance from S to ℓ: $\left|\left|\overrightarrow{SM}\right|\right| = \left|\left|\overrightarrow{PS}\right|\right| \sin\theta$ $= \frac{||\overrightarrow{PS}\times\mathbf{v}||}{||\mathbf{v}||} = \left|\left|\overrightarrow{PS} \times \hat{\mathbf{v}}\right|\right|$.



Example 2.4.4. *Find the distance from a point $S(1,3,2)$ to the line ℓ : $\mathbf{r} = (0,1,2) + t\,(1,0,1)\,,\ t \in \mathbb{R}$.*

Solution. Let P denote the point $(0,1,2)$ on ℓ, $\mathbf{v} = \mathbf{i}+\mathbf{k}$ (direction vector), and θ be the angle between $\overrightarrow{PS} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ and ℓ.

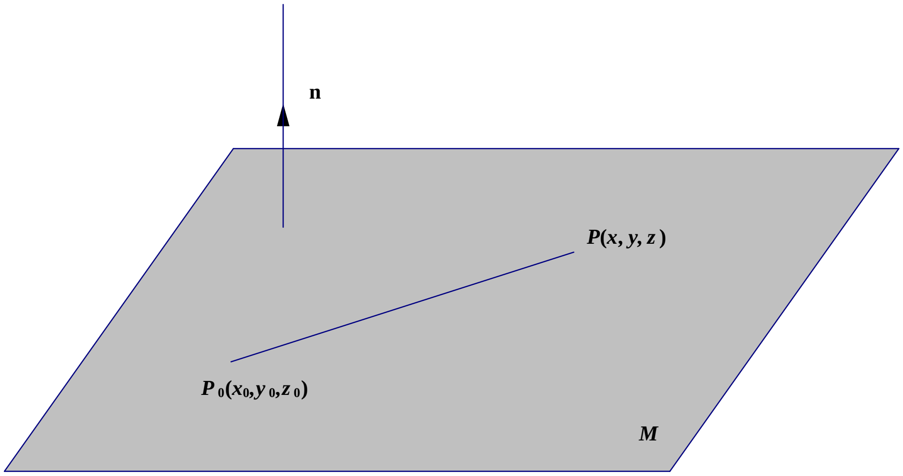
Then the distance from P to $\ell =$

$$\left|\left|\overrightarrow{PS}\right|\right| \sin\theta = \frac{\left|\overrightarrow{PS} \times \mathbf{v}\right|}{|\mathbf{v}|}$$

$$= \frac{\left|\begin{pmatrix} \\ \\ \\ \end{pmatrix} \times \begin{pmatrix} \\ \\ \\ \end{pmatrix}\right|}{||\mathbf{v}||} = \frac{\left|\begin{pmatrix} \\ \\ \\ \end{pmatrix}\right|}{\sqrt{2}} = \frac{3}{\sqrt{2}}\text{units.}$$

2.4.2 Plane

A plane in a space is determined by a point and its "inclination". This inclination can be specified by a vector $\mathbf{n}$ that is normal, or perpendicular, to the plane.



Suppose a plane M contains a point $P_0 = (x_0, y_0, z_0)$ and a nonzero vector $\mathbf{n} = (a, b, c)$ normal to the plane. If $P = (x, y, z)$ is any other point lying on this plane, then

P_0P and $\mathbf{n}$ are perpendicular.

So, we have

$$\mathbf{n}\cdot\overrightarrow{P_0P} = 0.$$

i.e.,

$$\mathbf{n} \cdot \overrightarrow{OP} = \underbrace{\mathbf{n} \cdot \overrightarrow{OP_0}}_{\text{constant}},$$

or simply:

$$\mathbf{r} \cdot \mathbf{n} = d.$$

This is called the **vector equation of the plane**. It follows from $\mathbf{n}\cdot\overrightarrow{P_0P} = 0$ that

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = 0$$

which can be expanded to

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

This is called the **scalar equation of the plane** through $P_0 = (x_0, y_0, z_0)$ with normal vector $\mathbf{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. It can be simplified to

$$ax + by + cz = d,$$

where $d = ax_0 + by_0 + cz_0$.

Example 2.4.5. *Find the scalar equation for the plane through* $P_0\,(1, 2, 3)$ *and perpendicular to* $\mathbf{n} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$.

Solution The required equation is

$$4\,(x - 1) + 5\,(y - 2) + 6\,(z - 3) = 0.$$

Simplifying, we have

$$4x + 5y + 6z = 32.$$

Example 2.4.6. *Find the scalar equation of the plane passing through* $A\,(1, 0, 0)$, $B\,(0, 1, 0)$ *and* $C\,(0, 0, 1)$.

Solution We first determine a vector normal to the plane. The vector $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are vectors parallel to the plane. Therefore their cross product is a vector normal to the plane. Let $\mathbf{n} = \overrightarrow{AB} \times \overrightarrow{AC}$.

$$\mathbf{n} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \\ \\ \\ \end{pmatrix}.$$

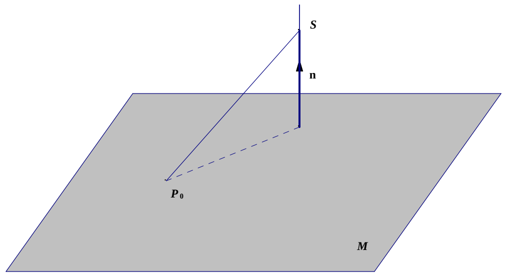
The scalar equation of the plane is

$$1 \cdot (x - 1) + 1 \cdot (y - 0) + 1 \cdot (z - 0) = 0,$$

or

$$x + y + z = 1.$$

Distance from a point to a plane



If P is a point of the plane with normal $\mathbf{n}$, then the distance from any point S to the plane is the length of the vector projection of $\overrightarrow{PS}$ to $\mathbf{n}$, which is equal to

$$= \left|\frac{\overrightarrow{PS}\cdot\mathbf{n}}{||\mathbf{n}||}\right| = \left|\overrightarrow{PS}\cdot\hat{\mathbf{n}}\right|.$$

Example 2.4.7. *Find the distance from the point $S\,(1,0,-1)$ to the plane $x-2y+z=1$.*

Solution

A normal vector of the plane $x-2y+z=1$ is $\mathbf{n}=(1,-2,1)$, with $\|\mathbf{n}\|=\sqrt{6}$.

We find a point P on the given plane by finding (x,y,z) which satisfies the equation $x-2y+z=1$.

Let $x=0$ and $y=0$, we have $z=$ ____________.

Thus, we may take $P=(0,0,1)$.

The required distance is given by $\left|\overrightarrow{PS}\cdot\hat{\mathbf{n}}\right|$ which is

$$\left|((1,0,-1)-(0,0,1))\cdot\frac{1}{\sqrt{6}}(1,-2,1)\right| = \frac{1}{\sqrt{6}}.$$

Angles between two planes

The angle between two planes is defined to be the (acute) angle between their respective normal vectors.

Example 2.4.8. *Find the angles between the planes $x-2y+2z=1$ and $6x-4y+3z=7$.*

Solution. The vectors $\mathbf{n}_1=\begin{pmatrix}1\\-2\\2\end{pmatrix}$ and $\mathbf{n}_2=\begin{pmatrix}6\\-4\\3\end{pmatrix}$ are the normal vectors of the respective planes.

If θ is the (acute) angle between the planes, then $\cos\theta=\left|\frac{\mathbf{n}_1\cdot\mathbf{n}_2}{||\mathbf{n}_1||||\mathbf{n}_2||}\right|$. Therefore

$$\theta=\cos^{-1}\left|\frac{\mathbf{n}_1\cdot\mathbf{n}_2}{|\mathbf{n}_1|\,|\mathbf{n}_2|}\right| = \cos^{-1}\frac{|6+8+6|}{\sqrt{1+4+4}\sqrt{36+16+9}} = \cos^{-1}\left(\rule{3cm}{0.4pt}\right).$$

Chapter 3

Matrices

A matrix is simply a rectangular array of numbers or other mathematical objects. Rectangular arrays of real numbers (or complex numbers) arise in many contexts and in sciences and engineering as well as social sciences. Beside providing a neat representation of data from which we can obtain further information, matrices have wide applications in this modern generation.

3.1 Matrix Notation and Terminology

For positive integers m and n, an $m \times n$ **matrix** A is a rectangular array of mn numbers (real or complex numbers) arranged in m horizontal rows and n vertical columns:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}.$$

The (i, j)th entry (or simply ij-entry) is the term a_{ij} found in the ith row and jth column. The ith row of A, where $1 \leq i \leq m$, is

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix},$$

while the jth column, for $1 \leq j \leq n$, is

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Notation and Terminology

1. Capital letters $A, B, C, \ldots$ are used to denote matrices, and lowercase letters to denote numerical quantities. Some examples:

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 9 & 6 & 3 \end{bmatrix}, T = \begin{pmatrix} a & b & c \\ d & e & f \\ x & y & z \end{pmatrix}, Z = \begin{bmatrix} z_1 & z_2 & z_3 & z_4 \\ y_1 & y_2 & y_3 & y_4 \end{bmatrix}$$

Both square brackets or round brackets are used to enclose the array of entries.

2. The number m of rows and the number n of columns describe the **size** of a matrix. We write it as $m \times n$, and read as 'm by n'.

3. The matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

is sometimes written as

$$[a_{ij}]_{m\times n} \text{ or simply } [a_{ij}].$$

To refer to the (i, j)th-entry of the matrix A, we may use the notation A_{ij}. So, $A_{ij} = a_{ij}$.

For $A = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 9 & 6 & 3 & 0 \\ -2 & -4 & -8 & \sqrt{2} \end{bmatrix}$, we have $A_{11} = 1, A_{21} = 9$ and $A_{34} =??$.

4. Usually, we match the letter denoting a matrix with the letter denoting its entries. For a matrix B, its ij-entry is b_{ij}.

5. When $m = 1$, the matrix has only one row,i.e.

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_n \end{bmatrix}$$

We call this matrix a **row matrix** (also known as **row vector**) .

6. When $n = 1$, the matrix has only one column,i.e.

$$B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}$$

We call such matrix a **column matrix** (also known as **column vector**) .

Of course, for a row matrix, we may simplify the entry using only one index, i.e. $C = [c_1 \ \ c_2 \ \ \cdots \ \ c_n]$.

7. The $m \times n$ matrix with zeros as its entries is called the **zero matrix** and we denote it by 0.

8. When $m = n$, we call A a **square matrix** of size n. (The rectangular array now looks like a square.)

   For an $n \times n$ square matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{nn} \end{bmatrix}.$$

   The entries $a_{11}, a_{22}, a_{33}, \ldots, a_{nn}$ are called the **diagonal entries**. They are on the main diagonal of A.

9. The $n \times n$ square matrix where all the entries along the diagonal from the top left to the bottom right are 1, and 0 elsewhere, is called the **identity matrix**. It is often denoted as I_n. E.g.,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

10. The $n \times n$ square matrix where all the off-diagonal entries (i.e. entries below and above the main diagonal) are 0 is called an **diagonal matrix**.

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & g \end{pmatrix} \quad B = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi \end{pmatrix}.$$

11. The $n \times n$ square matrix where all the entries below the main diagonal are 0 is called an **upper triangular matrix**.

$$A = \begin{pmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & g \end{pmatrix} \quad B = \begin{pmatrix} 1 & 3 & 5 & 0 \\ 0 & 3 & -2 & 9 \\ 0 & 0 & 0 & \pi \\ 0 & 0 & 0 & 0.8 \end{pmatrix}.$$

   In a similar way, a **lower triangular matrix** is a square matrix where all the entries above the main diagonal are 0. E.g.

$$C = \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \quad B = \begin{pmatrix} 53 & 0 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 & 0 \\ 1/2 & 0 & \pi & 0 & 0 \\ 0 & -1 & 0.4 & \sqrt{2} & 0 \\ 1 & 3 & 5 & 7 & 9 \end{pmatrix}.$$

12. Lastly, the term **scalars** refer to real numbers or complex numbers in discussing matrices (and vectors).

Definition 3.1.1. *Two matrices are defined to be* ***equal*** *if they have the* ***same size*** *and their corresponding entries are equal.*

Matrices A and B are equal if they have the same size (same number of rows and same number of columns) and $A_{ij} = B_{ij}$ for all i and j. (Here, note that if the size of both matrix is $m \times n$, then $1 \le i \le m$ and $1 \le j \le n$.)

Example 3.1.2. *Solve the following matrix equation for a, b, c and d.*

$$\begin{bmatrix} a-b & b+c \\ 3d+c & 2a-4d \end{bmatrix} = \begin{bmatrix} 8 & 1 \\ 7 & 6 \end{bmatrix}$$

Solution Note that both matrices are 2 by 2. For matrices to be equal, each corresponding entries must be equal. Therefore we have

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

Solving for a, b, c and d (Exercise), we obtain $a = 5, b = -3, c = 4, d = 1$.

Example 3.1.3. *Let A be a 3×4 matrix whose (i, j)-th entry is defined by $(A)_{ij} = (-1)^{i+j} 2i + j$. Then*

$$A = \begin{bmatrix} 3 & 0 & 5 & 2 \\ & 6 & & \\ & & 9 & \\ & & & \end{bmatrix}.$$

Example 3.1.4. *Let $B = [b_{ij}]$ be a 3×3 matrix where* $b_{ij} = \begin{cases} i+j & \textit{if } i > j \\ 0 & \textit{if } i = j \\ -j & \textit{if } i < j \end{cases}$

Find B.
Solution

$$B = \begin{bmatrix} 0 & & \\ & 0 & \\ & & 0 \\ & & \end{bmatrix}.$$

3.2 Arithmetic Operations of Matrices

We will study the arithmetic operations for matrices.

3.2.1 Addition & Subtraction

Definition 3.2.1. *If A and B are two $m \times n$ matrices, then*

(a) the ***addition*** *(or* ***sum****) $A + B$ is the matrix obtained by adding entries in the same positions, i.e.,*

$$(A+B)_{ij} = (A)_{ij} + (B)_{ij}.$$

(b) the ***difference*** *$A - B$ is the matrix obtained by subtracting entries of B from the corresponding entries of B, i.e.,*

$$(A-B)_{ij} = (A)_{ij} - (B)_{ij}.$$

WARNING Matrices of different sizes cannot be added or subtracted.

Example 3.2.2. *Let* $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ *and* $B = \begin{pmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}$. *Then*

$$A + B = \begin{pmatrix} 1+7 & 2+8 & 3+9 \\ 4+10 & 5+11 & 6+12 \end{pmatrix} = \begin{pmatrix} 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix}$$

$$A - B = \begin{pmatrix} 1-7 & 2-8 & 3-9 \\ 4-10 & 5-11 & 6-12 \end{pmatrix} = \begin{pmatrix} -6 & -6 & -6 \\ -6 & -6 & -6 \end{pmatrix}.$$

3.2.2 Scalar Multiplication

Here, by a scalar, we refer to a real number or a complex number.

Definition 3.2.3. *If α is a scalar and A is an $m \times n$ matrix, then the* ***scalar multiple*** *αA is the $m \times n$ matrix obtained by multiplying each entry of A by α, i.e.,*

$$(\alpha A)_{ij} = \alpha (A)_{ij}.$$

Example 3.2.4.

Let $A = \begin{pmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \end{pmatrix}$. *Then*

$$2A = \begin{pmatrix} 2 & 6 & 10 \\ 14 & 18 & 22 \end{pmatrix}, (-3)A = \begin{pmatrix} -3 & -9 & -15 \\ -21 & -27 & -33 \end{pmatrix}, \text{ and } \frac{1}{3}A = \begin{pmatrix} \frac{1}{3} & 1 & \frac{5}{3} \\ \frac{7}{3} & 3 & \frac{11}{3} \end{pmatrix},$$

Note that for $m \times n$ matrices A and B , the difference $A - B = A + (-1)B$. It is common practice to denote $(-1)B$ by $-B$.

Example 3.2.5. *Let* $A = \begin{pmatrix} 4 & 3 & 2 \\ 1 & 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 7 & 2 & 0 \\ -5 & 3 & -1 \end{pmatrix}$ *and* $C = \begin{pmatrix} 3 & -6 & 9 \\ 3 & 0 & 12 \end{pmatrix}$. *Then*

$$2A - B + \frac{1}{3}C = \begin{pmatrix} 2 & 2 & 7 \\ 8 & 3 & 7 \end{pmatrix}$$

.

For matrices, when we perform arithmetic operations like addition, difference and scalar multiplication, we are basically performing similar arithmetic operations sum, difference and multiplication on numbers on 'entry'-level. Thus, we would expect properties such as associativity, commutativity and distributive to hold for such matrix operations.

Theorem 3.2.6. *Assuming the sizes of matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid.*

1. $A + B = B + A$ *(Commutative Law for addition)*
2. $(A + B) + C = A + (B + C)$ *(Associative Law for addition)*
3. $A + 0 = 0 + A = A$ *(Additive Identity)*
4. $A + (-A) = 0$ *(Additive Inverse)*
5. $0 - A = -A$
6. $\alpha(A \pm B) = \alpha A \pm \alpha B$
7. $(\alpha \pm \beta)A = \alpha A \pm \beta A$
8. $\alpha(\beta A) = (\alpha\beta)A$

3.2.3 Matrix Multiplication

Can we multiply two matrices in the similar way i.e. by multiplying corresponding entries like what we have done for addition, subtraction or scalar multiplication?

It turns out that such a definition is not very helpful for most problems. Experience has led mathematicians to the following more useful definition of matrix multiplication.

Definition 3.2.7. *If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then the* ***matrix product*** *(or simply the* ***product****) AB is the $m \times n$ matrix whose (i, j)th entry is determined by*

$$(AB)_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{ir}B_{rj} = \sum_{k=1}^{r} A_{ik}B_{kj}.$$

Note that the (i, j)th entry of AB is the value obtained by taking the dot product of the vector formed by the ith row of A and that formed by the jth column of B.

$$AB = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & & \vdots \\ A_{i1} & A_{i2} & \cdots & A_{ir} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mr} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1j} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2j} & \cdots & B_{2n} \\ \vdots & \vdots & & \vdots & & \\ B_{r1} & B_{r2} & \cdots & B_{rj} & \cdots & B_{rn} \end{pmatrix}.$$

Example 3.2.8.

(a) Let $A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 1 & 4 \end{pmatrix}$ *and* $B = \begin{pmatrix} -2 \\ 4 \\ 2 \end{pmatrix}$. *Note that A is 2×3 matrix and B is 3×1, the product AB will be a 2×1 matrix.*

To find the entries of AB:

$(1, 1)$*-th entry:*

$(2, 1)$*-th entry:*

$(3, 1)$*-th entry:*

Thus, we have $AB = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$.

(b) *Let* $A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 1 & 4 \end{pmatrix}$ *and* $C = \begin{pmatrix} -3 \\ 0 \\ 2 \\ \sqrt{3} \end{pmatrix}$.

Note that A *is* 2×3 *matrix and* C *is* 4×1*, the product* AC *is not defined.*

(c) *Let* $A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 1 & 4 \end{pmatrix}$ *and* $D = \begin{pmatrix} -2 & 0 \\ 4 & 5 \\ 2 & 1 \end{pmatrix}$.

Find AD *and* DA*. Is* AA *defined?*

Solution

$$AD = \begin{pmatrix} 4 & \quad \\ 6 & \quad \end{pmatrix}, DA = \begin{pmatrix} -2 & -4 & 2 \\ & & 16 \\ & & \end{pmatrix}.$$

Remark

1. Matrix multiplication is not commutative.

2. It is not true that if $AB = 0$, then $A = 0$ or $B = 0$. (Exercise: Find two non-zero 2×2 matrices A and B with $AB = 0$.)

Without computing the entire product, we may compute a particular row or column of a matrix product AB as follows:

Proposition 3.2.9 (Optional)**.** *Let* A *and* B *be* $m \times r$ *and* $r \times n$ *matrices respectively. Then*

(a) *the* j*th column of* (AB) *is the matrix product of* A *and the* j*th column of* B*;*

$$j\textit{th column of } (AB) = A[j\textit{th column of } B].$$

(b) *the* i*th row of* (AB) *is the matrix product of the* i*th row of* A *and the matrix* B*;*

$$i\textit{th row of } (AB) = [i\textit{th row of } A]B.$$

Example 3.2.10. *Let* $A = \begin{pmatrix} 4 & 3 & 2 \\ 1 & 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \\ 0 & -3 & -2 \end{pmatrix}$.

Find (a) the 2nd column of AB *and (b) the last row of* AB.

Solution

Identity Matrix

Recall that the identity matrix I_n is the square matrix I_n of size n with (i, j)th-entry

$$(I_n)_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j \end{cases}$$

Thus,

$$I_1 = \begin{pmatrix} 1 \end{pmatrix}, I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ldots$$

The role of identity matrices in matrix multiplication is like the number 1 in usual multiplication. Some algebraic properties on matrix multiplication are recorded in the next theorem.

Theorem 3.2.11. *Assuming the sizes of matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid.*

1. $AI_n = A$ *and* $I_m A = A$ *if* A *is* $m \times n$. *(Identity)*
2. $(AB)C = A(BC)$ *(Associative Law for multiplication);*
3. $A(B + C) = AB + AC$ *(Left distributive law)*
4. $(A + B)C = AC + BC$; *(Right distributive law)*
5. $A0 = 0, 0A = 0$

3.3 Transpose

Definition 3.3.1. *For an $m \times n$ matrix A, the matrix A^T obtained by interchanging the rows and columns of A is called the* ***transpose*** *of A. Thus, the (i, j)th-entry of A^T is*

$$(A^T)_{ij} = (A)_{ji}$$

Example 3.3.2. *Let* $A = \begin{pmatrix} 4 & 3 & 2 \\ 1 & 3 & 1 \end{pmatrix}$ *and* $B = \begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \\ 0 & -3 & -2 \end{pmatrix}$. *Then*

$$A^T = \begin{pmatrix} \quad & \quad & \\ & & \\ & & \end{pmatrix}, B^T = \begin{pmatrix} \quad & \quad & \quad \\ & & \\ & & \end{pmatrix}$$

Theorem 3.3.3. *If the sizes of the matrices are such that the stated operations can be performed, then*

Proposition 3.3.4. *(a)* $(A^T)^T = A$

(b) $(A \pm B)^T = A^T \pm B^T$

(c) $(\alpha A)^T = \alpha(A^T)$, *where α is a scalar.*

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

Proof. (Exercise.)

3.4 Matrix Inverse

In this section, we shall discuss matrix invertibility with respect to matrix multiplication. We shall discuss uniqueness, properties of matrix inverses, finding inverses.

3.4.1 Invertible Matrices

Definition 3.4.1. *Let A be a $n \times n$ square matrix. If there is another square matrix B such that $AB = I_n$ and $BA = I_n$, then A is said to be* ***invertible*** *(or* ***non-singular****), and B is called an* ***inverse*** *of A. If no such matrix B can be found, then A is said to be* ***not invertible*** *(or* ***singular****).*

Example 3.4.2. *Let* $A = \begin{pmatrix} -1 & -2 \\ 3 & 5 \end{pmatrix}$ *and* $B = \begin{pmatrix} 5 & 2 \\ -3 & -1 \end{pmatrix}$.

Note that $AB = I_2$ and $BA = I_2$ (Verify these as an exercise).

Since $AB = I_2$ and $BA = I_2$, we conclude that A is invertible, by definition.

Example 3.4.3. *Let* $C = \begin{pmatrix} 0 & 0 & 0 \\ a & b & c \\ d & e & f \end{pmatrix}$. *Is there a matrix D such that $CD = I$?*

Solution

Example 3.4.4. *Let* $E = \begin{pmatrix} a & b & c \\ 2a & 2b & 2c \\ d & e & f \end{pmatrix}$. *Is there a matrix F such that $EF = I$?*

Solution

Note that matrix invertibility is defined for square matrices only. It is clear that zero square matrices are singular. By Proposition 3.2.9, we deduce that the following are singular.

Proposition 3.4.5.

(a) A square matrix with a row (or a column) of zeros is singular.

(b) A square matrix with a row (or a column) which is a multiple of another row (or column) is singular. In particular, a square matrix with identical rows (or columns) is singular.

A row R_i is a **linear combination** of other rows R_j, where $j \neq i$, means that there are scalars α_j's such that

$$R_i = \sum_{j \neq i} \alpha_j R_j = \alpha_1 R_1 + \alpha_2 R_2 + \cdots + \alpha_{i-1} R_{i-1} + \alpha_{i+1} R_{i+1} + \cdots \alpha_n R_2 n.$$

Proposition 3.4.6. *A square matrix in which a row (or a column) is a linear combination of other rows (or columns) is singular.*

Example 3.4.7. *Consider* $G = \begin{pmatrix} a & b & c \\ d & e & f \\ 2a - 5d & 2b - 5e\& \; 2c - 5f \end{pmatrix}$.

Note that $R_3 = 2R_1 + (-5)R_2$. *The matrix* G *singular.*

3.4.2 The inverse A^{-1}

Now we prove that an invertible matrix cannot have more than one inverses. In other words, if an inverse exists, it is unique.

Proposition 3.4.8. *If* B *and* $\widehat{B}$ *are both inverses of* A, *then* $B = \widehat{B}$.

Proof. Note that
$$AB = I\&BA = I, \text{ and } A\widehat{B} = I\&\widehat{B}A = I.$$

□

The above proposition says that the inverse of an invertible matrix A is unique. In view of this, we shall denote the inverse of A by A^{-1}. Thus, we have

$$AA^{-1} = I \text{ and } A^{-1}A = I.$$

It follows immediately from the above equation and definition of matrix invertibility that the inverse A^{-1} of a matrix A is invertible.

Proposition 3.4.9. *If* A *is an invertible matrix, then the inverse* A^{-1} *is invertible and*

$$\left(A^{-1}\right)^{-1} = A.$$

3.4.3 Invertible 2 by 2 matrices

For a 2×2 matrix, there is a good characterization of invertibility Moreover, there is a nice formula for its inverse when it is invertible. This is recorded in the following proposition.

Proposition 3.4.10. *The* 2×2 *matrix* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *is invertible if and only if* $ad - bc \neq 0$. *In this case the inverse* A^{-1} *is given by the formula*

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{3.1}$$

Proof.

$\Longleftarrow$: Suppose $ad - bc \neq 0$. We verify (Exercise.) that the matrix equations are satisfied

$$\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} A = I_2 \ \&\ A\,\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = I_2.$$

By the definition of matrix invertibility, A is invertible and its inverse A^{-1} is

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

$\Longrightarrow$: It remains to prove that if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible, then $ad - bc \neq 0$.

We prove the contrapositive statement: If $ad - bc = 0$, then the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is not invertible.

Suppose $ad - bc = 0$. We shall consider two cases : (a) $a = 0$ (b) $a \neq 0$.

For the case (a) where $a = 0$, note that either $b = 0$ or $c = 0$. If $b = 0$, then A has a row of zero and this it is not invertible. If $c = 0$, then the first row and second row of A are multiple of each other. Hence A is not invertible.

For The case (b) where $a \neq 0$, we have $d = \frac{bc}{a}$. This the second row of A is a scalar multiple of the fist row of A. ($\frac{c}{a}$ of the first row) Thus, A is not invertible.

Therefore, if $ad - bc = 0$, then A is not invertible. Equivalently, if A is invertible, then $ad - bc \neq 0$.

□

Remark 3.4.11. *The above result provides a useful characterization of invertible* 2×2 *matrices. For a* 2×2 *matrix* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, *the value* $ad - bc$ *is called the* ***determinant*** *of* A. *The above result also provides a formula for the inverse* A^{-1} *if the matrix* A *is invertible.*

Example 3.4.12. *Determine whether each of the following 2×2 matrices is invertible. If it is, find its inverse.*

(a) $A = \begin{pmatrix} 5 & 3 \\ 7 & 9 \end{pmatrix}$

(b) $B = \begin{pmatrix} -8 & 4 \\ 6 & -3 \end{pmatrix}$

(c) $C = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$.

Remark 3.4.13.

(a) *For matrices of other sizes, there is similar result on invertibility via determinant which will be discussed in latter section. There is also a formula for the inverse of an invertible matrix. However, the formula is more complicated than the case for 2×2 matrices.*

(b) *We may use Gaussian method to determine whether a square matrix is invertible and also to find its inverse. This will be dealt with in another mathematics Course.*

However, for a diagonal matrix, there is an easy way to determine its invertibility and its inverse.

Example 3.4.14. *Find the inverse of the following diagonal matrix, if $abcd \neq 0$.*

$$\begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix}$$

More generally, we have the following result for diagonal matrices.

Proposition 3.4.15. *Suppose $A = [a_{ij}]$ is a diagonal matrix. Then A is invertible if and only of its diagonal entries are non-zero, i.e., $a_{ii} \neq 0$ for every i. When A is invertible, A^{-1} is a diagonal matrix whose ith diagonal entry is $\dfrac{1}{a_{ii}}$ for each i.*

Proposition 3.4.16 (Optional)**.** *Suppose A a triangular matrix. Then A is invertible if and only if all diagonal entries are non zero. Moreover, the inverse of invertible upper (resp. lower) triangular matrix will be an upper (resp. lower) triangular matrix.*

(The inverse of a general invertible triangular matrix is quite messy to include here.)

Now, we study some properties of invertible matrices.

Proposition 3.4.17. *Let A be an invertible matrix. Then we have*

1. $AB = AC \Longrightarrow B = C$ *and*
2. $BA = CA \Longrightarrow B = C$.

Proposition 3.4.18. *Let A and B be invertible matrices. Then the matrix product AB is invertible and*

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

Proposition 3.4.19. *Suppose A is invertible.*

(a) Then its transpose A^T is invertible. In this case,

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T.$$

(b) For any nonzero scalar α, the matrix αA is invertible and

$$(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}.$$

3.5 Power

Definition 3.5.1. *Let A be a square matrix. We define the nonnegative integer powers of A to be*

$$A^0 = I \quad A^n = \underbrace{AA\cdots A}(n > 0).$$

Moreover, if A is invertible, then we define the negative integer powers to be

$$A^{-n} = \left(A^{-1}\right)^n = \underbrace{A^{-1}A^{-1}\cdots A^{-1}}(n > 0)$$

Parallel to real numbers, we have

$$A^rA^s = A^{r+s}, (A^r)^s = A^{rs}.$$

Example 3.5.2. *Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Find $A^2, A^3, A^4, A^{-2}, A^{-3}$.*

Solution

We also have the following laws of exponents.

Proposition 3.5.3. *Let n be an integer. If A be an invertible matrix, then A^n is invertible and*

$$(A^n)^{-1} = \left(A^{-1}\right)^n.$$

3.6 Determinants

Determinants via Cofactor Expansion

Recall that a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible if and only if $ad - bc \neq 0$. This special number $ad - bc$ is known as the determinant of the 2×2 square matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. It is denoted by the symbol $\det(A)$. The determinant is a function defined on square matrices. For a general $n \times n$ square matrix A, where $n \geq 3$, we shall compute the $\det(A)$ inductively via Cofactor Expansion.

3.6.1 Cofactors

Definition 3.6.1. *Suppose A is an $n \times n$ matrix.*

*(1) The (i, j)-**minor of** A is defined to be the determinant of the submatrix that remains after the ith-row and the jth-column are deleted from A. It is denoted by M_{ij}.*

*(2) The (i, j)-**cofactor of** A is the number $(-1)^{i+j} M_{ij}$. It is denoted by C_{ij}.*

Let M_{ij} be the determinant of the submatrix that remains after the ith-row and the jth-column are deleted from A.

*The (i, j)**th cofactor** of A is the number $(-1)^{i+j}M_{ij}$. It is denoted by C_{ij}.*

Note The value M_{ij} is called (i, j)th minor of A.

Example 3.6.2. *Consider the matrix*

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

(a) Find the $(1, 1)$th cofactor and $(2, 3)$th cofactor of A.

(b) Calculate C_{12} and C_{31}.

Solution

Determinant & Cofactors

Cofactors are used in the evaluation of determinants in an inductive way. First we note the determinants of matrices of sizes 1×1 and 2×2.

1. The determinant of a 1×1 matrix $[a]$ is a, i.e. $\det([a]) = a$.
2. The determinant of the 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $ad - bc$. We write it as

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc, \quad \text{or} \quad \det(A) = ad - bc,$$

   which can be obtained by computing the sum of the products on the rightward arrow and subtracting the products in the leftward arrow.

To compute $\det(A)$, where A is $n \times n$ with $n \geq 3$, we proceed as follows:

1. Select a row of A, say ith row.
2. Multiply each entry of the selected row by its cofactor, i.e., $a_{ik}C_{ik}$.
3. Add all the resulting products obtained in the last step, i.e., $\sum_{k=1}^{n} a_{ik}C_{ik}$. This number is $\det(A)$.

For an $n \times n$ matrix A, we have

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots a_{in}C_{in} = \sum_{j=1}^{n} a_{ij}Cij$$

(cofactor expansion along the ith-row)

Example 3.6.3. *Find the determinant of the matrix* $A = \begin{bmatrix} 1 & 5 & 0 \\ -3 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}$ *by cofactor expansion.*

[Solution] We evaluate the determinant of A along second row:

$$\det(A) = \begin{vmatrix} 1 & 5 & 0 \\ -3 & 2 & 1 \\ 1 & 2 & 1 \end{vmatrix} = -3\cdot(-1)^{2+1} \Big| \qquad \Big| + 2\cdot(-1)^{2+2} \Big| \qquad \Big| + 1\cdot(-1)^{2+3} \Big| \qquad \Big|$$

$$= \qquad = 20.$$

Note that using minors, we have

$$\det(A) = (-1)^{1+j}a_{1j}M_{1j} + (-1)^{2+j}a_{2j}M_{2j} + \cdots (-1)^{n+j}a_{nj}M_{nj} = (-1)^j \sum_{i=1}^{n} (-1)^i a_{ij}M_{ij},$$

in which the sign of each consecutive terms alternates, with j being held at constant.

The "checkerboard matrix" S is a matrix with entries $S_{ij} = (-1)^{i+j}$. The following are examples of 3 by 3 and 4 by 4 checkerboard matrices.

$$S = \begin{bmatrix} +1 & -1 & +1 \\ -1 & +1 & -1 \\ +1 & -1 & +1 \end{bmatrix} \text{ and } S = \begin{bmatrix} +1 & -1 & +1 & -1 \\ -1 & +1 & -1 & +1 \\ +1 & -1 & +1 & -1 \\ -1 & +1 & -1 & +1 \end{bmatrix}.$$

We may refer to the checkerboard matrix for the sign of $(-1)^{i+j}$ when computing determinants.

For instance, if we compute determinant along third row, the checkerboard gives us the signs $\{+, -, +\}$:

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 5 & 0 \\ -3 & 2 & 1 \\ 1 & 2 & 1 \end{vmatrix} = (1)\begin{vmatrix} 5 & 0 \\ 2 & 1 \end{vmatrix} - (2)\begin{vmatrix} 1 & 0 \\ -3 & 1 \end{vmatrix} + (1)\begin{vmatrix} 1 & 5 \\ -3 & 2 \end{vmatrix} \\ &= 5 - 2 + 17 = 20. \end{aligned}$$

Determinant of via Cofactors along Column

Theorem 3.6.4. *For a square matrix A, the determinants of A and its transpose A^T are the same, i.e.,* $\det(A) = \det(A^T)$.

Therefore, instead of performing cofactor expansion along a selected row, we may also evaluate the determinant of A by cofactor expansion along a selected column.

1. Select a column of A, say jth column. (So, we say that we perform cofactor expansion along the jth-column.)
2. Multiply each entry a_{kj} of the selected row by its corresponding cofactor C_{kj}, i.e., $a_{kj}C_{kj}$.
3. Add all the resulting products obtained in the last step gives us the determinant of A, i.e.,

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} = \sum_{k=1}^{n} a_{kj}C_{kj}.$$

(cofactor expansion along the jth-column)

Example 3.6.5. *Compute determinant using cofactor expansion along the second column.*

$$det(A) = \begin{vmatrix} 1 & 5 & 0 \\ -3 & 2 & 1 \\ 1 & 2 & 1 \end{vmatrix}$$

Solution Referring to the checkerboard, which gives us the signs $\{-, +, -\}$ along the second column:

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 5 & 0 \\ -3 & 2 & 1 \\ 1 & 2 & 1 \end{vmatrix} = -5 \begin{vmatrix} -3 & 1 \\ 1 & 1 \end{vmatrix} + 2 \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 1 & 0 \\ -3 & 1 \end{vmatrix} \\ &= (-5)(-4) + (2)(1) - 2(1) = 20. \end{aligned}$$

Determinant of 3×3 matrices

The determinant of the 3×3 matrix $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ is

$$\det(A) = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

Remark 3.6.6. *An easy way to help us in computing the determinant of 3×3 matrix (ONLY) is by recopying the first and second columns next to the third column of A and followed by computing the sum of the products on the rightward arrows and subtracting the products in the leftward arrows.*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

Example 3.6.7. *Evaluate the determinant using the above method:*

(a) $\begin{vmatrix} -2 & 1 & 4 \\ 3 & 5 & -7 \\ 1 & 6 & 2 \end{vmatrix}$

(Ans -65)

(b) $\begin{vmatrix} -2 & 7 & 6 \\ 5 & 1 & -2 \\ 3 & 8 & 4 \end{vmatrix}$

(Ans 0)

Example 3.6.8. *Find the determinant of the matrix* $B = \begin{bmatrix} 1 & 1 & 5 & 0 \\ 0 & -3 & 2 & 1 \\ 0 & 1 & 2 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix}$ *by cofactor expansion.*

(Ans 12).

In general, one strategy for evaluating a determinant by cofactor expansion is to expand along a row or column having the largest number of zeros.

Example 3.6.9. *Compute* $\begin{vmatrix} 3 & 1 & 2 & 2 \\ 1 & 0 & -2 & 1 \\ 1 & 0 & 5 & -1 \\ 2 & 0 & 0 & 1 \end{vmatrix}$.

(Ans -1*).*

3.6.2 Determinants of Special Matrices.

In this section, we discuss determinants of some types of matrices.

Proposition 3.6.10.

1. *Suppose A is $n \times n$ matrix which as a row of zero or a column of zeros. Then* $\det(A) = 0$.
2. *If A has two rows (or columns) such that one of which is a multiple of the other, then* $\det(A) = 0$.
3. *Suppose A is a triangular matrix. Then* $\det(A) =$ *product of diagonal entries of A. In particular if A is a diagonal matrix, then* $\det(A) =$ *product of diagonal entries of A.*

Example 3.6.11. *The determinants of the following matrices are zero.*

$$\begin{bmatrix} 4 & 9 & 8 & 5 \\ -5 & -2 & 0 & 7 \\ 0 & 0 & 0 & 0 \\ 3 & 4 & \pi & \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} 4 & 9 & 8 & 5 \\ -5 & -2 & 0 & 7 \\ 6 & 8 & 2\pi & 1 \\ 3 & 4 & \pi & \frac{1}{2} \end{bmatrix}.$$

Example 3.6.12. *Find the following determinants.*

(a) $\det(I_n) =$

(b) $\begin{vmatrix} 4 & 0 & 0 & 5 \\ 0 & -2 & 0 & 7 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{vmatrix}$ *(c)* $\begin{vmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{vmatrix}$

Theorem 3.6.13. *For two $n \times n$ matrices A and B,*

$$\det(AB) = \det(A)\det(B).$$

Proof. (Proof Omitted.) $\square$

Corollary 3.6.14. *If A is an $n \times n$ invertible matrix, then $\det(A) \neq 0$. Moreover,*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Proof. From $AA^{-1} = I$, we have

$$\det(AA^{-1}) = \det(I), \text{ i.e. }, \det(A)\det(A^{-1}) = 1.$$

Since $\det(A) \neq 0$, we have $\det(A^{-1}) = \dfrac{1}{\det(A)}$. $\square$

3.7 Adjoint of A (Optional)

In this section, we define an adjoint of a square matrix A, and obtain an important result relating A and its adjoint. From this result, we derive the equivalent statement for invertibility of a matrix A and nonzero determinant. It also gives a formula for the inverse of an invertible matrix.

Definition 3.7.1. *Let A be an $n \times n$ matrix, and C_{ij} be its (i,j)th cofactor. Then the matrix*

$$\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}$$

is called the ***matrix of cofactors from*** *A.*

Example 3.7.2. *Find the cofactor matrix of* $\begin{bmatrix} 1 & 5 & 0 \\ -3 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}$

Solution

(a) $C_{11} = (-1)^{1+1}\begin{vmatrix} 2 & 1 \\ 2 & 1 \end{vmatrix} = 0, C_{12} = (-1)^{1+2}\begin{vmatrix} -3 & 1 \\ 1 & 1 \end{vmatrix} = 4, C_{13} = (-1)^{1+3}\begin{vmatrix} -3 & 2 \\ 1 & 2 \end{vmatrix} = -8...$

So we have

$$C = \begin{bmatrix} 0 & 4 & -8 \\ -5 & 1 & 3 \\ 5 & -1 & 17 \end{bmatrix} \ \& \ \mathrm{adj}(A) = \begin{bmatrix} 0 & -5 & 5 \\ 4 & 1 & -1 \\ -8 & 3 & 17 \end{bmatrix}.$$

(b)

$$A\mathrm{adj}(A) = \begin{bmatrix} 20 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 20 \end{bmatrix} = 20\ I.$$

Definition 3.7.3. *Let A be an $n \times n$ matrix, and C_{ij} be the cofactor. The transpose of the matrix of cofactors is called the* ***adjoint of A*** *and is denoted by* adj(A).

That is,

$$(\mathrm{adj}(A))_{ij} = C_{ji}.$$

Example 3.7.4. *(a) Find the adjoint of* $A = \begin{bmatrix} 1 & 5 & 0 \\ -3 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}$. *(b) Evaluate $A(adj\,(A))$.*

The cofactor matrix

$$C = \begin{bmatrix} 0 & 4 & -8 \\ -5 & 1 & 3 \\ 5 & -1 & 17 \end{bmatrix}$$

$$\text{adj}(A) = C^T = \begin{bmatrix} 0 & -5 & 5 \\ 4 & 1 & -1 \\ -8 & 3 & 17 \end{bmatrix}.$$

(b)

$$A\text{adj}(A) = \begin{bmatrix} 20 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 20 \end{bmatrix} = 20\ I.$$

We state following relationship between adjoint, determinant and inverse of a matrix.

Theorem 3.7.5. *Let A be an $n \times n$ square matrix. Then*

$$A\ adj(A) = \det(A)I.$$

Proof. (Optional. A proof is included at the end of this chapter for students who are keen to read.) □

The next result follows from the above theorem.

Theorem 3.7.6. *Let A be an $n \times n$ square matrix. Then A is invertible if and only if* $\det(A) \neq 0$.

(Equivalently, the matrix A is singular if and only if $\det(A) = 0$.*)*

If A is invertible, then

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A).$$

Example 3.7.7. *Find the inverse of* $\begin{bmatrix} 1 & 5 & 0 \\ -3 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}$ *via adjoint.*

3.8 Cramer's Rule

A system of linear equations is a finite number of linear equations. For instance,

$$\begin{cases} 7x_1 & - & 2x_2 & + & 5x_3 & = & 3 \\ 3x_1 & + & x_2 & - & 4x_3 & = & -2 \end{cases}$$

which can be expressed as a matrix equation:

$$\begin{pmatrix} 7 & -2 & 5 \\ 3 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}.$$

For a linear system $\mathbf{Ax} = \mathbf{b}$ whose coefficient matrix A is invertible, there is a formula for its solution. The formula is known as Cramer's rule. It is useful for studying the mathematical properties of a solution without the need for solving the system.

Example 3.8.1. *The linear system*

$$\begin{array}{rcrcrcr} 2u & - & v & + & w & = & 3 \\ u & + & v & - & 3w & = & 5 \\ 5u & - & 4v & + & 9w & = & 4 \end{array}$$

is equivalent to

$$\begin{pmatrix} 2 & -1 & 1 \\ 1 & 1 & -3 \\ 5 & -4 & 9 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 4 \end{pmatrix}.$$

Theorem 3.8.2 (Cramer's Rule)**.** *If* $\mathbf{Ax} = \mathbf{b}$ *is a system of* n *linear equations in* n *unknowns such that* $\det(A) \neq 0$*, then the system has a unique solution, namely*

$$x_j = \frac{\det(A_j)}{\det(A)}, j = 1, 2, \ldots, n$$

where

$$A_j = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & b_1 & a_{1j+1} \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j-1} & b_2 & a_{2j+1} \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & b_n & a_{nj+1} \cdots & a_{nn} \end{bmatrix},$$

the matrix obtained by replacing the entries in the j*th column of* $\mathbf{A}$ *by the entries in the matrix*

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Example 3.8.3. *For each of the following linear system, determine whether Cramer's Rule is applicable. If it is determine, use it to solve the linear system. If it is not, use other method to solve the system.*

(a)

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

(b)

$$\begin{array}{rcrcr} 2a & + & 4b & = & 3 \\ 3a & + & 6b & = & 5 \end{array}$$

(c)

$$\begin{array}{rcrcrcr} 2u & - & v & + & w & = & 3 \\ u & + & v & - & 3w & = & 5 \\ 5u & - & 4v & + & 9w & = & 4 \end{array}$$

3.9 Proofs (Optional)

To prove Theorem 3.7.5, we need some result on cofactors.

In the evaluation of the determinant of a matrix A,we select the ith row (or jth column) of A and the ith row (or jth column) of its cofactor matrix, i.e.,

$$\det(A) = \sum_{k=1}^{n} a_{ik}C_{ik}.$$

What happens if we have selected two different rows, and compute the product of entries (of a selected row) and cofactors from different row(the other selected row)? That is, when $i \neq j$, what is the value of

$$\sum_{k=1}^{n} a_{ik}C_{jk}?$$

We shall see from the next example that the sum of such products will be zero.

Example 3.9.1. *Consider the matrix 3×3 matrix $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.*

Consider cofactors of entries along second row:

$$C_{21} = -(a_{12}a_{33} - a_{13}a_{32}), C_{22} = a_{11}a_{33} - a_{13}a_{31}, C_{23} = -(a_{11}a_{32} - a_{12}a_{31})$$

and entries along first row of A. Now, we compute

$$a_{11}C_{21} + a_{12}C_{22} + a_{13}C_{23}.$$

This is actually the determinant of the following matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

whose determinant is zero.

Proposition 3.9.2. *Let A be an $n \times n$ matrix. Suppose $i \neq j$. Then*

$$\sum_{k=1}^{n} a_{ik}C_{jk} = 0.$$

Proof. (Optional.) Suppose $i \neq j$. Let A' be the matrix which has the same row as A except the jth row. The jth row of A' is the ith row of A. That is, A' has two identical rows, namely ith row and jth row.

Then we have

$$\sum_{k=1}^{n} a_{ik}C_{jk} = \det(A') = 0.$$

□

Proof. Proof of Theorem 3.7.5 (Optional.)

It follows from

$$A\text{adj}(A) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{j1} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{j2} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{jn} & \cdots & C_{nn} \end{bmatrix}$$

that the ij-th entry of the product $A\text{adj}(A)$ is

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn}$$

Case $i = j$: The ij-th entry of the product $A\text{adj}(A)$ is $\det(A)$.

Case $i \neq j$: The ij-th entry of the product $A\text{adj}(A)$ is 0, since the entries and cofactors come from different rows.

In conclusion, we have $A\ adj(A) = \det(A)I$. □

Proof of Cramer's Rule

Proof. Since $\det(A) \neq 0$, the matrix A is invertible and $A^{-1} = \frac{1}{det(A)} adj(A)$.

Thus, the unique solution of $\mathbf{Ax} = \mathbf{b}$ is given by

$$\mathbf{x} = \frac{1}{det(A)}\text{adj}(A)\mathbf{b}.$$

Hence, we have $x_j = \frac{1}{det(A)}(j\text{th-row of adj}(A))\mathbf{b}$.

One can verify that $(j\text{th-row of adj}(A))\mathbf{b}$ is the determinant of A_j where

$$A_j = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & b_1 & a_{1j+1} \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j-1} & b_2 & a_{2j+1} \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & b_n & a_{nj+1} \cdots & a_{nn} \end{bmatrix},$$

the matrix obtained by replacing the entries in the jth column of $\mathbf{A}$ by the entries in the matrix

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

□