Math 2: Linear Algebra

FOR THE ELECTRONICS AND TELECOMMUNICATION STUDENTS

Lecture Notes selected and prepared by:
Andrzej Maćkiewicz
Technical University of Poznań

Abstract

Contents

Preface

These are lecture notes for a first course in linear algebra; the prerequisite is a good course in Precalculus. I personally believe that many more people need linear algebra than calculus, thus the material in these notes is absolutely fundamental for all modern engineers in the digitalization era.

Linear algebra is one of the core topics studied at university level by students on many different types of degree programme. Alongside calculus, it provides the framework for mathematical modelling in many diverse areas. This text sets out to introduce and explain linear algebra to students from electronics and telecommunication. It covers all the material that would be expected to be in most first-year university courses in the subject, together with some more advanced material that would normally be taught later.

This text represents our best effort at distilling from my experience what it is that I think works best in helping students not only to do linear algebra, but to understand it. I regard understanding as essential.

I have attempted to write a user-friendly, fairly interactive and helpful text, and I intend that it could be useful not only as a course text, but for self-study. These notes are quite informal, but they have been carefully read and criticized by the students, and their comments and suggestions have been incorporated. Although I've tried to be careful, there are undoubtedly some errors remaining. If you find any, please let me know.

Carefully designed examples and exercises are provided in the supplementary volume of "Linear Algebra , Problems , Solutions and Tips"

Andrzej Maćkiewicz
Poznań, September 2014

1
Complex numbers

Many practical mathematical problems, especially in physics and electronics, involve square roots of negative numbers (that is, complex numbers). For example, modern theories of heat transfer, fluid flow, damped harmonic oscillation, alternating current circuit theory, quantum mechanics, and relativity—all beyond the scope of this text—depend on the use of complex quantities. Therefore, our next goal is to extend our knowledge to the realm of complex numbers.

One excellent reason for generalizing to the complex number system is that we can take advantage of the *Fundamental Theorem of Algebra*, which states that every $n-$th degree polynomial can be factored completely when complex roots are permitted. Later we will see how this permits us to find additional (non-real) solutions to eigenvalue problems.

1.1 Are complex numbers necessary?

Much of mathematics is concerned with various kinds of equations, of which equations with numerical solutions are the most elementary. The most fundamental set of numbers is the set $\mathbb{N} = \{1, 2, 3, ...\}$ of *natural numbers*. If a and b are natural numbers, then the equation $x + a = b$ has a solution within the set of natural numbers if and only if $a < b$. If $a \geq b$ we must extend the number system to the larger set $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, 3...\}$ of *integers*. Here we get a bonus, for the equation $x + a = b$ has a solution $x = b - a$ in $\mathbb{Z}$ for all a and b in $\mathbb{Z}$.

If $a, b \in \mathbb{Z}$ and $a \neq 0$, then the equation $ax + b = 0$ has a solution in $\mathbb{Z}$ if and only if a divides b. Otherwise we must once again extend the number system to the larger set $\mathbb{Q}$ of *rational numbers*. Once again we get a bonus, for the equation $ax + b = 0$ has a solution $x = -b/a$ in $\mathbb{Q}$ for all $a \neq 0$ in $\mathbb{Q}$ and all b in $\mathbb{Q}$.

When we come to consider a *quadratic equation* $ax^2 + bx + c = 0$ (where $a, b, c \in \mathbb{Q}$ and $a \neq 0$) we encounter our first real difficulty. We may safely assume that a, b and c are integers: if not , we simply multiply the equation by a suitable positive integer. The standard solution to the equation is given by the familiar formula

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

Let us denote $b^2 - 4ac$, the *discriminant of the equation*, by Δ. If Δ is the square of an integer (what is often called a *perfect square*) then the equation has rational solutions, and if Δ is positive then the two solutions are in the extended set $\mathbb{R}$ of real numbers. But if $\Delta < 0$ then there is no solution even within $\mathbb{R}$,

We have already carried out three extensions (to $\mathbb{Z}$, to $\mathbb{Q}$, to $\mathbb{R}$) from our starting point in natural numbers, and there is no reason to stop here. We can modify the standard formula to obtain

$$x = \frac{-b \pm \sqrt{(-1)\,(4ac - b^2)}}{2a}$$

where $4ac - b^2 > 0$. If we postulate the existence of $\sqrt{(-1)}$, then we get a "solution"

$$x = \frac{-b \pm \sqrt{(-1)}\sqrt{(4ac - b^2)}}{2a}.$$

Of course we know that there is no real number $\sqrt{(-1)}$, but the idea seems in a way to work. If we look at a specific example,

$$x^2 + 4x + 13 = 0,$$

and decide to write i for $\sqrt{(-1)}$, the formula gives us two solutions $x = -2+3i$ and $x = -2-3i$. If we use normal algebraic rules, replacing i^2 by -1 whenever it appears, we find that

$$\begin{aligned}(-2+3i)^2 + 4(-2+3i) + 13 &= (-2)^2 + 2(-2)(3i) + (3i)^2 - 8 + 12i + 13 \\ &= 4 - 12i - 9 - 8 + 12i + 13 \quad (\text{since } i^2 = -1) \\ &= 0.\end{aligned}$$

and the validity of the other root can be verified in the same way. We can certainly agree that if there is a number system containing "numbers" $a + bi$, where $a, b \in \mathbb{R}$, then they will add and multiply according to the rules

$$(a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + a_2) + (b_1 + b_2)i \tag{1.1}$$

$$(a_1 + b_1 i)(a_2 + b_2 i) = (a_1 a_2 - b_1 b_2) + (a_1 b_2 + b_1 a_2)i. \tag{1.2}$$

We shall see shortly that there is a way, closely analogous to our picture of real numbers as points on a line, of visualizing these new complex numbers.

Can we find equations that require us to extend our new complex number system (which we denote by $\mathbb{C}$) still further? No, in fact we cannot: the important *Fundamental Theorem of Algebra*, (whose proof is beyond the scope of this text), states that, for all $n \geq 1$, every polynomial equation

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

with coefficients $a_0, a_1, ..., a_n$ in $\mathbb{C}$ and $a_n \neq 0$, has all its roots within $\mathbb{C}$. This is one of many reasons why the number system $\mathbb{C}$ is of the highest importance in the development and application of mathematical ideas.

1.2 Sums and products

Complex numbers can be defined as ordered pairs (x, y) of real numbers that are to be interpreted as points in the *complex plane*, with rectangular coordinates x and y, just as real numbers x are thought of as points on the real line. When real numbers x are displayed as points $(x, 0)$ on the real axis, it is clear that the set of complex numbers includes the real numbers as a subset. Complex numbers of the form $(0, y)$ correspond to points on the y axis and are called *pure imaginary numbers* when $y \neq 0$. The y axis is then referred to as the *imaginary axis*.

It is customary to denote a complex number (x, y) by z, so that (see Fig. 1.1)

$$z = (x, y). \tag{1.3}$$

The real numbers x and y are, moreover, known as the *real* and *imaginary parts* of z, respectively; and we write

$$x = \operatorname{Re} z, \quad y = \operatorname{Im} z. \tag{1.4}$$

Two complex numbers z_1 and z_2 are *equal* whenever they have the same real parts and the same imaginary parts. Thus the statement $z_1 = z_2$ means that z_1 and z_2 correspond to the same point in the complex, or z, plane.

The *sum* $z_1 + z_2$ and *product* $z_1 z_2$ of two complex numbers $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$ are defined as follows:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \tag{1.5}$$

$$(x_1, y_1)(x_2, y_2) = (x_1 x_2 - y_1 y_2, y_1 x_2 + x_1 y_2). \tag{1.6}$$

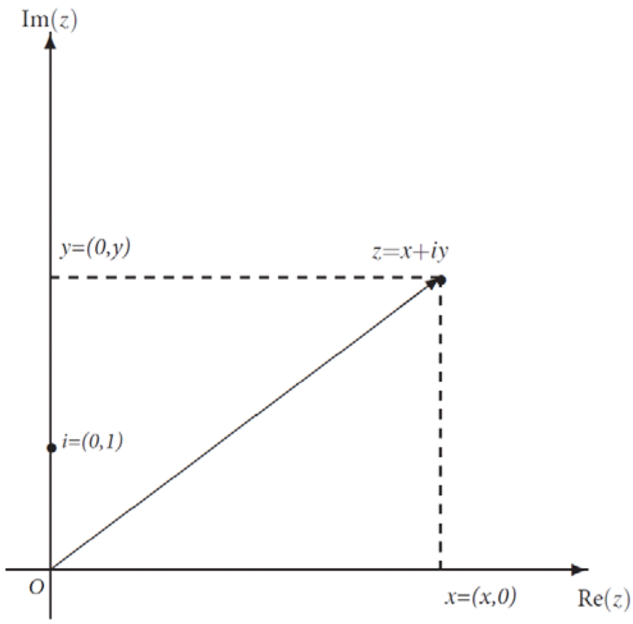


Fig. 1.1. Complex number z as a point in the complex plane.

Note that the operations defined by equations (1.5) and (1.6) become the usual operations of addition and multiplication when restricted to the real numbers:

$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0),$$

$$(x_1, 0)(x_2, 0) = (x_1 x_2, 0).$$

The complex number system is, therefore, a natural extension of the real number system. Any complex number $z = (x, y)$ can be written

$$z = (x, 0) + (0, y),$$

and it is easy to see that $(0, 1)(y, 0) = (0, y)$. Hence

$$z = (x, 0) + (0, 1)(y, 0);$$

and if we think of a real number as either x or $(x, 0)$ and let i denote the pure imaginary number $(0, 1)$, as shown in (Fig. 1.1), it is clear that[1]

$$z = x + iy. \tag{1.7}$$

[1] In electrical engineering, the letter j is used instead of i.

Also, with the convention that $z^2 = zz$, $z^3 = z^2 z$, etc., we have

$$i^2 = (0,1)(0,1) = (-1,0),$$

or

$$i^2 = -1. \tag{1.8}$$

Because $(x,y) = x + iy$, definitions (1.5) and (1.6) become

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i\,(y_1 + y_2)\,, \tag{1.9}$$

$$(x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i\,(y_1x_2 + x_1y_2)\,. \tag{1.10}$$

Observe that the right-hand sides of these equations can be obtained by formally manipulating the terms on the left as if they involved only real numbers and by replacing i^2 by -1 when it occurs. Also, observe how equation (1.10) tells us that any complex number times zero is zero. More precisely,

$$z{\cdot}0 = (x + iy)(0 + i0) = 0 + i0 = 0$$

for any $z = x + iy$.

1.3 Basic algebraic properties

Various properties of addition and multiplication of complex numbers are the same as for real numbers. We list here the more basic of these algebraic properties and verify some of them. Most of the others are verified in the exercises.

The *commutative laws*

$$z_1 + z_2 = z_2 + z_1 \tag{1.11}$$

and the *associative laws*

$$z_2z_1 = z_2z_1 \tag{1.12}$$

follow easily from the definitions in Sec. 1.2 of addition and multiplication of complex numbers and the fact that real numbers obey these laws. For example, if

$$z_1 = (x_1, y_1) \quad \text{and} \quad z_2 = (x_2, y_2),$$

then

$$z_1 + z_2 = (x_1 + x_2, y_1 + y_2) = (x_2 + x_1, y_2 + y_1) = z_2 + z_1.$$

Verification of the rest of the above laws, as well as the *distributive law*

$$z\,(z_1 + z_2) = zz_1 + zz_2 \tag{1.13}$$

is similar.

According to the commutative law for multiplication, $iy = yi$. Hence one can write $z = x+yi$ instead of $z = x+iy$. Also, because of the associative laws, a sum $z_1 + z_2 + z_3$ or a product $z_1z_2z_3$ is well defined without parentheses, as is the case with real numbers.

The additive identity $0 = (0, 0)$ and the multiplicative identity $1 = (1, 0)$ for real numbers carry over to the entire complex number system. That is,

$$z + 0 = z \quad \text{and} \quad z \cdot 1 = z \tag{1.14}$$

for every complex number z. Furthermore, 0 and 1 are the only complex numbers with such properties (see Exercise **??**).

There is associated with each complex number $z = (x, y)$ an *additive inverse*

$$-z = (-x, -y) \tag{1.15}$$

satisfying the equation $z + (-z) = 0$. Moreover, there is only one additive inverse for any given z, since the equation

$$(x, y) + (u, v) = (0, 0)$$

implies that

$$u = -x \quad \text{and} \quad v = -y.$$

For any nonzero complex number $z = (x, y)$, there is a number z^{-1} such that $zz^{-1} = 1$. This *multiplicative inverse* is less obvious than the additive one. To find it, we seek real numbers u and v, expressed in terms of x and y, such that

$$(x, y)(u, v) = (1, 0).$$

According to equation (1.6), Sec. 1.2, which defines the product of two complex numbers, u and v must satisfy the pair

$$xu - yv = 1, \qquad yu + xv = 0$$

of linear simultaneous equations; and simple computation yields the unique solution

$$u = \frac{x}{x^2 + y^2}, \qquad v = \frac{-y}{x^2 + y^2}.$$

So the multiplicative inverse of $z = (x, y)$ is

$$z^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right) \quad \text{for} \quad z \neq 0. \tag{1.16}$$

The inverse z^{-1} is not defined when $z = 0$. In fact, $z = 0$ means that $x^2 + y^2 = 0$; and this is not permitted in expression (1.16).

1.4 Further properties

In this section, we mention a number of other algebraic properties of addition and multiplication of complex numbers that follow from the ones already described. Inasmuch as such properties continue to be anticipated because they also apply to real numbers, the reader can easily pass to Sec. 1.5 without serious disruption.

We begin with the observation that the existence of multiplicative inverses enables us to show that *if a product z_1z_2 is zero, then so is at least one of the factors z_1 and z_2*. For suppose that $z_1z_2 = 0$ and $z_1 \neq 0$. The inverse z^{-1} exists; and any complex number times zero is zero (Sec. 1.2). Hence

$$z_2 = z_2{\cdot}1 = z_2(z_1z^{-1}) = (z^{-1}z_1)z_2 = z^{-1}(z_1z_2) = z^{-1}{\cdot}0 = 0.$$

That is, if $z_1z_2 = 0$, either $z_1 = 0$ or $z_2 = 0$; or possibly both of the numbers z_1 and z_2 are zero. Another way to state this result is that *if two complex numbers z_1 and z_2 are nonzero, then so is their product z_1z_2.*

Subtraction and division are defined in terms of additive and multiplicative inverses:

$$z_1 - z_2 = z_1 + (-z_2), \tag{1.17}$$

$$\frac{z_1}{z_2} = z_1z_2^{-1}, \quad \text{where} \quad z_2 \neq 0. \tag{1.18}$$

Thus, in view of expressions (1.15) and (1.16) in Sec. 1.3,

$$z_1 - z_2 = (x_1, y_1) + (-x_2, -y_2) = (x_1 - x_2, y_1 - y_2) \tag{1.19}$$

and

$$\begin{aligned} \frac{z_1}{z_2} &= (x_1, y_1)\left(\frac{x_2}{x_2^2 + y_2^2}, \frac{-y_2}{x_2^2 + y_2^2}\right) \\ &= \left(\frac{x_1x_2 - y_1y_2}{x_2^2 + y_2^2}, \frac{y_1x_2 + x_1y_2}{x_2^2 + y_2^2}\right) \quad \text{if} \quad z_2 \neq 0, \end{aligned} \tag{1.20}$$

and when $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$.

Using $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, one can write expressions (1.19) and (1.20) here as

$$z_1 - z_2 = (x_1 - x_2) + i\,(y_1 - y_2) \tag{1.21}$$

and

$$\frac{z_1}{z_2} = \frac{x_1x_2 - y_1y_2}{x_2^2 + y_2^2} + i\frac{y_1x_2 + x_1y_2}{x_2^2 + y_2^2} \quad \text{if} \quad z_2 \neq 0. \tag{1.22}$$

Although expression (1.22) is not easy to remember, it can be obtained by writing

$$\frac{z_1}{z_2} = \frac{(x_1 + iy_1)\,(x_2 - iy_2)}{(x_2 + iy_2)\,(x_2 - iy_2)}, \tag{1.23}$$

multiplying out the products in the numerator and denominator on the right, and then using the property

$$\frac{z_1 + z_2}{z_3} = (z_1 + z_2)\, z_3^{-1} = z_1 z_3^{-1} + z_2 z_3^{-1} = \frac{z_1}{z_3} + \frac{z_2}{z_3} \quad (z_3 \neq 0). \tag{1.24}$$

The motivation for starting with equation (1.23) appears in Sec. 1.6.

Example 1.1 *The method is illustrated below:*

$$\frac{4+i}{2-3i} = \frac{(4+i)\,(2+3i)}{(2-3i)\,(2+3i)} = \frac{5+14i}{13} = \frac{5}{13} + \frac{14}{13}i.$$

There are some expected properties involving quotients that follow from the relation

$$\frac{1}{z_2} = z_2^{-1} \quad (z_2 \neq 0). \tag{1.25}$$

which is equation (1.18) when $z_1 = 1$. Relation (1.25) enables us, for instance, to write equation (1.18) in the form

$$\frac{z_1}{z_2} = z_1 \left(\frac{1}{z_2}\right) \quad (z_2 \neq 0)\,.$$

Also, by observing that (see Exercise **??**)

$$(z_1 z_2)(z_1^{-1} z_2^{-1}) = (z_1 z_1^{-1})(z_2 z_2^{-1}) = 1 \quad (z_1 \neq 0,\ z_2 \neq 0),$$

and hence that $z_1^{-1} z_2^{-1} = (z_1 z_2)^{-1}$, one can use relation (1.25) to show that

$$\left(\frac{1}{z_1}\right)\left(\frac{1}{z_2}\right) = z_1^{-1} z_2^{-1} = (z_1 z_2)^{-1} = \frac{1}{z_1 z_2} \quad (z_1 \neq 0,\ z_2 \neq 0). \tag{1.26}$$

Another useful property, to be derived in the exercises, is

$$\left(\frac{z_1}{z_3}\right)\left(\frac{z_2}{z_4}\right) = \frac{z_1 z_2}{z_3 z_4} \quad (z_3 \neq 0,\ z_4 \neq 0). \tag{1.27}$$

Example 1.2 *Computations such as the following are now justified:*

$$\begin{aligned}\left(\frac{1}{2-3i}\right)\left(\frac{1}{1+i}\right) &= \frac{1}{(2-3i)(1+i)} = \frac{1}{5-i}\cdot\frac{5+i}{5+i}\\ &= \frac{5+i}{(5-i)(5+i)} = \frac{5+i}{26} = \frac{5}{26}+\frac{i}{26}\\ &= \frac{5}{26}+\frac{1}{26}i.\end{aligned}$$

Finally, we note that the *binomial formula* involving real numbers remains valid with complex numbers. That is, if z_1 and z_2 are any two nonzero complex numbers, then

$$(z_1+z_2)^n = \sum_{k=0}^{n}\binom{n}{k} z_1^k z_2^{n-k} \quad (n=1,2,...,) \tag{1.28}$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (k=0,1,2,...,n)$$

and where it is agreed that $0! = 1$. The proof, by mathematical induction, is the same as in the real case.

1.5 Vectors and moduli

It is natural to associate any nonzero complex number $z = x + iy$ with the directed line segment, or vector, from the origin to the point (x, y) that represents z in the complex plane. In fact, we often refer to z as the point $z$or the vector z. In Fig. 1.2the numbers $z_1 = 1 + 2i$ and $z_2 = 3 + i$ are displayed graphically as both points and radius vectors.

Generally, there are two geometric interpretations of the complex number $z = x + iy$:

1. as the point (x, y) in the xy-plane,
2. as the vector from the origin to (x, y).

In each representation the x-axis is called the *real axis* and the y-axis is called the *imaginary axis.* Both representations are *Argand diagram* for $x+iy$.

When $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, the sum

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

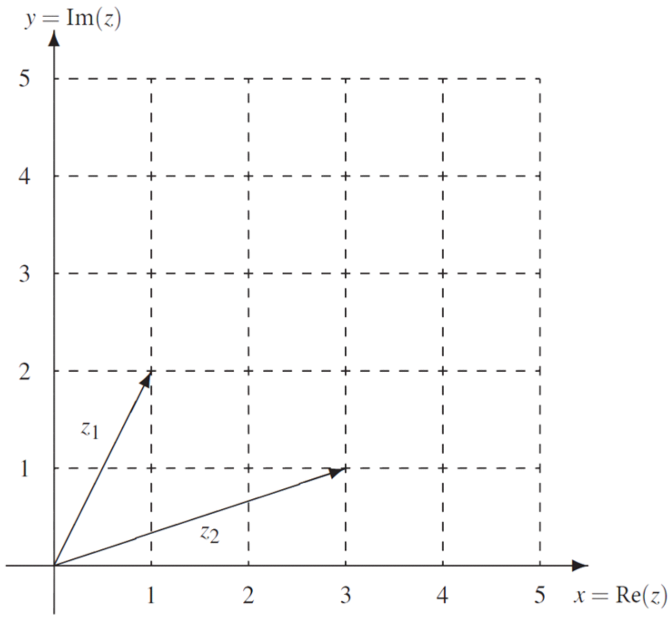


Fig. 1.2. The numbers $z_1 = 1 + 2i$ and $z_2 = 3 + i$ displayed graphically.

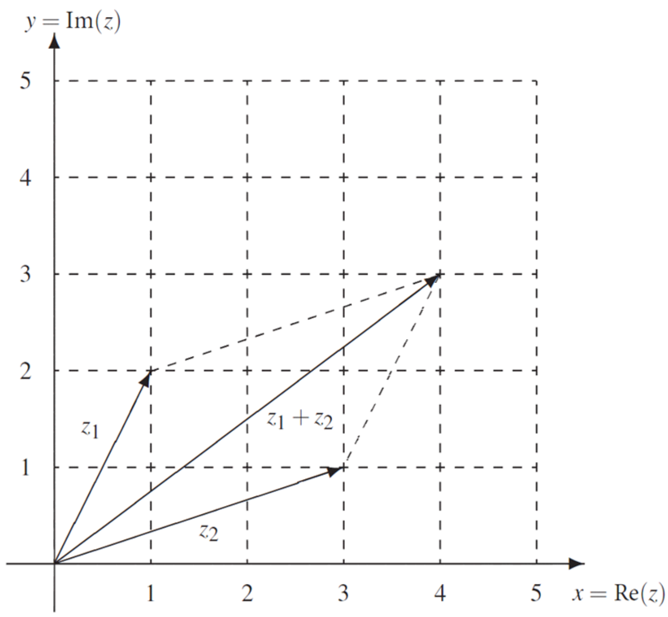


Fig. 1.3. Sum of two complex numbers obtained vectorially.

corresponds to the point (x_1+x_2, y_1+y_2). It also corresponds to a vector with those coordinates as its components. Hence z_1+z_2 may be obtained vectorially as shown in Fig. 1.3.

Although the product of two complex numbers z_1 and z_2 is itself a complex number represented by a vector, that vector lies in the same plane as the vectors for z_1 and z_2. Evidently, then, this product is neither the scalar nor the vector product used in ordinary vector analysis[2].

The vector interpretation of complex numbers is especially helpful in extending the concept of absolute values of real numbers to the complex plane. The modulus, or absolute value, of a complex number $z = x + iy$ is defined as the nonnegative real number $\sqrt{x^2 + y^2}$ and is denoted by $|z|$; that is,

$$|z| = \sqrt{x^2 + y^2}. \tag{1.29}$$

Geometrically, the number $|z|$ is the distance between the point (x, y) and the origin, or the length of the radius vector representing z. It reduces to the usual absolute value in the real number system when $y = 0$. Note that while the inequality $z_1 < z_2$ is meaningless unless both z_1 and z_2 are real, the statement $|z_1| < |z_2|$ means that the point z_1 is closer to the origin than the point z_2 is.

Example 1.3 *Since $|1 + 2i| = \sqrt{5}$ and $|3 + i| = \sqrt{10}$ we know that the point $z_1 = 1 + 2i$ is closer to the origin than $z_2 = 3 + 1$ is. Here we were able to establish this fact algebraically, without examining the Fig. 1.2.*

The distance between two points (x_1, y_1) and (x_2, y_2) is $|z_1 - z_2|$. This is clear from Fig. 1.2, since $|z_1 - z_2|$ is the length of the vector representing the number

$$z_1 - z_2 = z_1 + (-z_2);$$

and, by translating the radius vector $z_1 - z_2$, one can interpret $z_1 - z_2$ as the directed line segment from the point (x_2, y_2) to the point (x_1, y_1). Alternatively, it follows from the expression

$$z_1 - z_2 = (x_1 - x_2) + i(y_1 - y_2)$$

and definition (1.29) that

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

[2] These products will be defined later in this book.

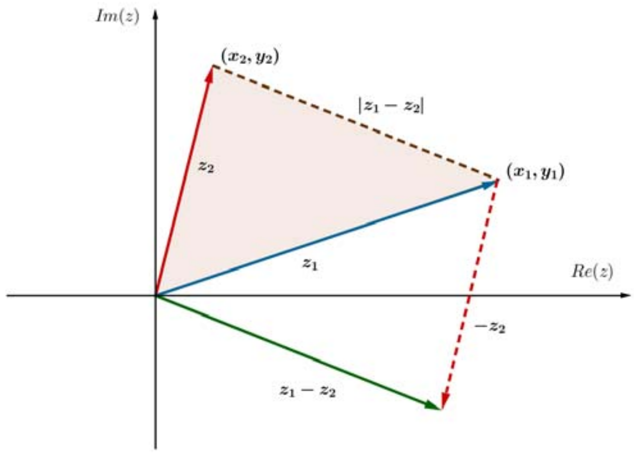


Fig. 1.4. Distance between two complex numbers.

The complex numbers z corresponding to the points lying on the circle with center z_0 and radius R thus satisfy the equation $|z - z_0| = R$, and conversely. We refer to this set of points simply as the circle $|z - z_0| = R$ or $\{z : |z - z_0| = R\}$.

Example 1.4 *The equation $|z-1+3i| = 2$ represents the circle whose center is $z_0 = (1, -3)$ and whose radius is $R = 2$.*

It also follows from definition (1.29) that the real numbers $|z|$, $\operatorname{Re} z = x$, and $\operatorname{Im} z = y$ are related by the equation

$$|z|^2 = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2 . \tag{1.30}$$

Thus

$$\operatorname{Re} z \leq |\operatorname{Re} z| \leq |z| \quad \text{and} \quad \operatorname{Im} z \leq |\operatorname{Im} z| \leq |z| . \tag{1.31}$$

We turn now to the *triangle inequality* , which provides an upper bound for the modulus of the sum of two complex numbers z_1 and z_2

$$|z_1 + z_2| \leq |z_1| + |z_2| . \tag{1.32}$$

This important inequality is geometrically evident in Fig. 1.3, since it is merely a statement that the length of one side of a triangle is less than or equal to the sum of the lengths of the other two sides. We can also see from

Fig. 1.3 that inequality (1.32) is actually an equality when 0, z_1, and z_2 are collinear. Another, strictly algebraic, derivation is given below.

Proof. (**Triangle inequality**) Let $z_1 = a + ib$ and $z_2 = c + id$. Then $0 \leq (ad - bc)^2 = a^2d^2 - 2abcd + b^2c^2$, so

$$2abcd \leq a^2d^2 + b^2c^2$$
$$a^2c^2 + 2abcd + b^2d^2 \leq a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2$$
$$(ac + bd)^2 \leq \left(a^2 + b^2\right)\left(c^2 + d^2\right)$$
$$2(ac + bd) \leq 2\sqrt{a^2 + b^2}\sqrt{c^2 + d^2}$$
$$a^2 + 2ac + c^2 + b^2 + 2bd + d^2 \leq a^2 + b^2 + 2\sqrt{a^2 + b^2}\sqrt{c^2 + d^2} + c^2 + d^2$$
$$(a + c)^2 + (b + d)^2 \leq \left(\sqrt{a^2 + b^2} + \sqrt{c^2 + d^2}\right)^2$$
$$|z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2$$
$$|z_1 + z_2| \leq |z_1| + |z_2|$$

■

An immediate consequence of the triangle inequality is the fact that

$$|z_1 + z_2| \geq ||z_1| - |z_2|| \tag{1.33}$$

To derive inequality (1.33), we write

$$|z_1| = |(z_1 + z_2) + (-z_2)| \leq |z_1 + z_2| + |-z_2|,$$

which means that

$$|z_1 + z_2| \geq |z_1| - |z_2|\,. \tag{1.34}$$

This is inequality (1.33) when $|z_1| \geq |z_2|$. If $|z_1| < |z_2|$, we need only interchange z_1 and z_2 in inequality (1.34) to arrive at

$$|z_1 + z_2| \geq -(|z_1| - |z_2|),$$

which is the desired result. Inequality (1.33) tells us, of course, that the length of one side of a triangle is greater than or equal to the difference of the lengths of the other two sides.

Because $|-z_2| = |z_2|$, one can replace z_2 by $-z_2$ in inequalities (1.32) and (1.33) to summarize these results in a particularly useful form:

$$|z_1 \pm z_2| \leq |z_1| + |z_2| \tag{1.35}$$

$$|z_1 \pm z_2| \geq |z_1| - |z_2|\,. \tag{1.36}$$

When combined, inequalities (1.35) and (1.36) become

$$||z_1| - |z_2|| \leq |z_1 \pm z_2| \leq |z_1| + |z_2|\,. \tag{1.37}$$

Example 1.5 *If a point z lies on the unit circle $|z| = 1$ about the origin, it follows from inequalities (1.35) and (1.36) that*

$$|z - 2| \leq |z| + 2 = 3$$

and

$$|z - 2| \geq ||z| - 2| = 1.$$

The triangle inequality (1.32) can be generalized by means of mathematical induction to sums involving any finite number of terms:

$$|z_1 + z_2 + \cdots + z_n| \leq |z_1| + |z_2| + ... + |z_n| \quad (n = 2, 3, ...). \tag{1.38}$$

To give details of the induction proof here, we note that when $n = 2$, inequality (1.38) is just inequality (1.32). Furthermore, if inequality (1.38) is assumed to be valid when $n = m$, it must also hold when $n = m + 1$ since, by inequality (1.32),

$$\begin{aligned} |(z_1 + z_2 + \cdots + z_m) + z_{m+1}| &\leq |z_1 + z_2 + \cdots + z_m| + |z_{m+1}| \\ &\leq (|z_1| + |z_2| + \cdots + |z_m|) + |z_{m+1}|. \end{aligned}$$

1.6 Complex conjugates

The *complex conjugate*, or simply the *conjugate*, of a complex number $z = x + iy$ is defined as the complex number $x - iy$ and is denoted by $\bar{z}$; that is,

$$\bar{z} = x - iy. \tag{1.39}$$

The number $\bar{z}$ is represented by the point $(x, -y)$, which is the reflection in the real axis of the point (x, y) representing z (Fig. 1.5). Note that

$$\overline{(\bar{z})} = z \quad \text{and} \quad |\bar{z}| = |z|$$

for all z.

If $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, then

$$\overline{z_1 + z_2} = (x_1 + x_2) - i(y_1 + y_2) = (x_1 - iy_1) + (x_2 - iy_2).$$

So the conjugate of the sum is the sum of the conjugates:

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2 \tag{1.40}$$

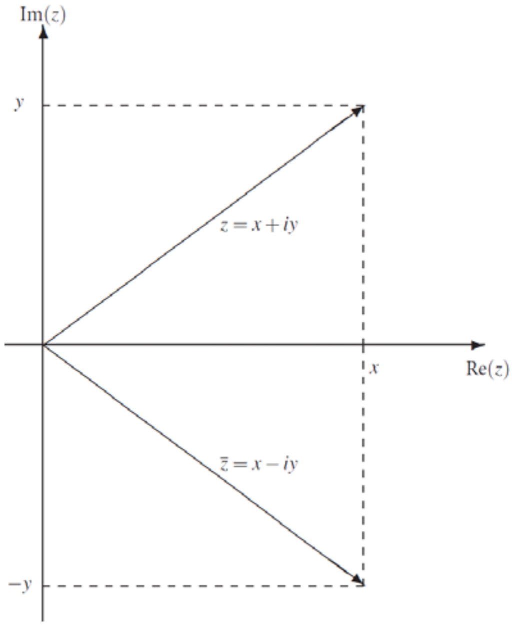


Fig. 1.5. The conjugate of a complex number $z = x + iy$.

In like manner, it is easy to show that

$$\overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2, \tag{1.41}$$

$$\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2 \tag{1.42}$$

and

$$\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}. \tag{1.43}$$

The sum $z + \bar{z}$ of a complex number$z = x + iy$ and its conjugate $z = x - iy$ is the real number $2x$, and the difference $z - \bar{z}$ is the pure imaginary number $2iy$. Hence

$$\operatorname{Re} z = \frac{z + \bar{z}}{2} \quad \text{and} \quad \operatorname{Im} z = \frac{z - \bar{z}}{2i}. \tag{1.44}$$

An important identity relating the conjugate of a complex number $z = x + iy$ to its modulus is

$$z\bar{z} = |z|^2 \tag{1.45}$$

where each side is equal to $x^2 + y^2$. It suggests the method for determining a quotient z_1/z_2 that begins with expression (1.22), Sec. 1.4. That method is, of course, based on multiplying both the numerator and the denominator of z_1/z_2 by $\bar{z}_2$, so that the denominator becomes the real number $|z_2|^2$.

Example 1.6 *As an illustration,*

$$\frac{-1+3i}{2-i} = \frac{(-1+3i)\,(2+i)}{(2-i)\,(2+i)} = \frac{-5+5i}{|2-i|^2} = -1+i.$$

See also the example in Sec. 1.4.

Identity (1.45) is especially useful in obtaining properties of moduli from properties of conjugates noted above. We mention that

$$|z_1 z_2| = |z_1|\,|z_2| \tag{1.46}$$

and

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|} \quad (z_2 \neq 0) \tag{1.47}$$

Property (1.46) can be established by writing

$$|z_1 z_2|^2 = (z_1 z_2)(\overline{z_1 z_2}) = (z_1 z_2)(\bar{z}_1 \bar{z}_2) = (z_1 \bar{z}_1)(z_2 \bar{z}_2) = |z_1|^2 |z_2|^2 = (|z_1||z_2|)^2$$

and recalling that a modulus is never negative. Property (1.47) can be verified in a similar way.

Example 1.7 *Property (1.46) tells us that* $|z^2| = |z|^2$ *and* $|z^3| = |z|^3$*. Hence if* z *is a point inside the circle centered at the origin with radius* 2*, so that* $|z| < 2$*, it follows from the generalized triangle inequality (1.38) in Sec. 1.5 that*

$$|z^3 + 3z^2 - 2z + 1| \leq |z|^3 + 3|z|^2 + 2|z| + 1 < 25.$$

Remark 1.8 *We have seen that the triangle inequality* $|z_1+z_2| \leq |z_1|+|z_2|$ *indicates that the length of the vector* z_1+z_2 *cannot exceed the sum of the lengths of the individual vectors* z_1 *and* z_2*.But the results given in (1.46 and 1.47) are interesting.The product* $z_1 z_2$ *and quotient* z_1/z_2*,* $(z_2 \neq 0)$*, are complex numbers and so are vectors in the complex plane.The equalities* $|z_1 z_2| = |z_1||z_2|$ *and* $|z_1/z_2| = |z_1|/|z_2|$ *indicate that the lengths of the vectors* $z_1 z_2$ *and* z_1/z_2 *are exactly equal to the product of the lengths and to the quotient of the lengths, respectively, of the individual vectors* z_1 *and* z_2*.*

Example 1.9 (An upper bound) *Find an upper bound for* $\left|\dfrac{-1}{z^4 - 5z + 1}\right|$ *if* $|z| = 2$.

Solution: By the previous result, the absolute value of a quotient is the quotient of the absolute values.Th us with $|-1| = 1$, we want to find a positive real number M such that

$$\frac{1}{|z^4 - 5z + 1|} \leq M.$$

To accomplish this task we want the denominator as small as possible.By (1.33) we can write

$$\begin{aligned}\left|z^4-(5z+1)\right| &\geq \left|\left|z^4\right|-|5z+1|\right| \geq \left|\left|z^4\right|-(5\,|z|+1)\right| \\ &= \left|\left|z^4\right|-5\,|z|-1\right| = |16-10-1| = 5.\end{aligned}$$

Hence for $|z|=2$ we have

$$\left|\frac{-1}{z^4-5z+1}\right| \leq \frac{1}{5}. \qquad \square$$

1.7 Polar coordinate system

So far, you have been representing graphs of equations as collections of points on the rectangular coordinate system, where and represent the directed distances from the coordinate axes to the point. In this section, we introduce a new system for assigning coordinates to points in the plane *polar coordinates.* We start with an origin point, called the *pole*, and a ray called the *polar axis.* We then locate a point P using two coordinates, $(r;\theta)$, where r represents a *directed distance*[3] from the pole and is a measure of rotation from the polar axis (see Fig. 1.6). Roughly speaking, the polar coordinates $(r;\theta)$ of a point measure 'how far out' the point is from the pole (that's r), and 'how far to rotate' from the polar axis, (that's θ). For example, if we wished to plot the point P with polar coordinates $(4,\frac{5\pi}{6})$, we'd start at the pole, move out along the polar axis 4 units, then rotate $\frac{5\pi}{6}$ radians counter-clockwise.

The standard table of cosine and sine values can be used to generate the following figure (1.7), which should be committed to memory and will be useful next. In rectangular coordinates, each point has a unique representation. This is not true for polar coordinates. For instance, the coordinates $(r;\theta)$ and $(r;\theta+2\pi)$ represent the same point. Another way to obtain multiple representations of a point is to use negative values for r. If $r<0$, we begin by moving in the opposite direction on the polar axis from the pole. As you may have guessed, $\theta<0$ means the rotation away from the polar axis is clockwise instead of counter-clockwise. Because r is a *directed distance*, the coordinates $(r;\theta)$ and $(-r;\theta+\pi)$ represent the same point. In general, the point can be represented as

$$(r;\theta) = (r;\ \theta \pm 2n\pi) \quad \text{or} \quad (r;\theta) = (-r;\ \theta \pm (2n+1)\pi)$$

[3] We will explain more about this momentarily.

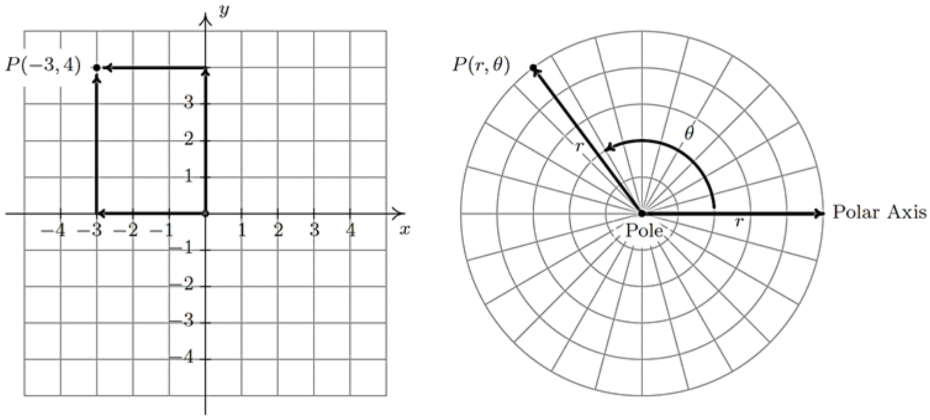


Fig. 1.6. Cartesian and polar coordinates.

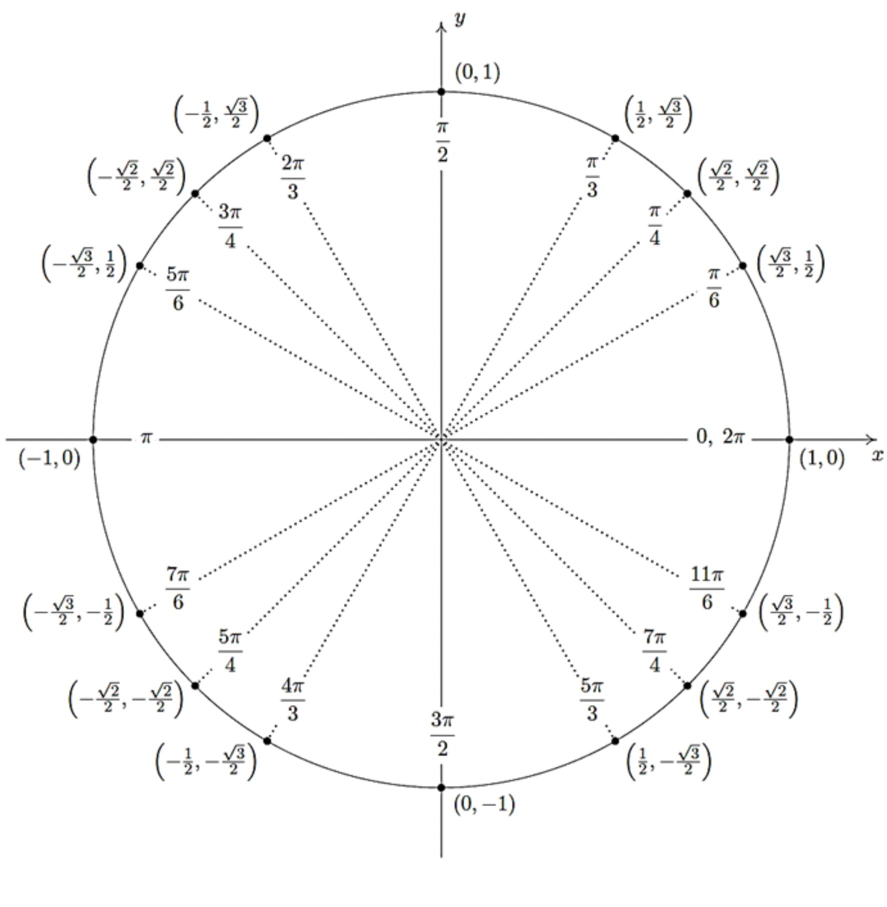


Fig. 1.7. Important points on the unit circle

where n is any integer. Moreover, the pole is represented by $(0,\theta)$ where θ is any angle.

Example 1.10 *The point $(3,-3\pi/4)$ has three additional polar representation with $-2\pi<\theta<2\pi$*

$$\begin{array}{lcll}\left(3,-\frac{3\pi}{4}+2\pi\right) & = & \left(3,\frac{5\pi}{4}\right) & \textit{Add } 2\pi \textit{ to } \theta.\\ \left(-3,-\frac{3\pi}{4}-\pi\right) & = & \left(-3,-\frac{7\pi}{4}\right) & \textit{Replace } r \textit{ by } -r, \textit{ subtract } \pi \textit{ from } \theta.\\ \left(-3,-\frac{3\pi}{4}+\pi\right) & = & \left(-3,\frac{\pi}{4}\right) & \textit{Replace } r \textit{ by } -r, \textit{ add } \pi \textit{ to } \theta.\end{array}$$

Next, we marry the polar coordinate system with the Cartesian (rectangular) coordinate system. To do so, we identify the pole and polar axis in the polar system to the origin and positive x-axis, respectively, in the rectangular system. We get the following result.

Theorem 1.11 *(Conversion Between Rectangular and Polar Coordinates): Suppose P is represented in rectangular coordinates as $(x;y)$ and in polar coordinates as (r,θ). Then*

- *Polar-to-Rectangular:* $x=r\cos(\theta)$ *and* $y=r\sin(\theta)$,
- *Rectangular-to-Polar:* $x^2+y^2=r^2$ *and* $\tan(\theta)=\frac{y}{x}$ *(provided $x\neq 0$).*

Proof. We know from elementary trigonometry, that if $Q(x;y)$ is the point on the terminal side of an angle, plotted in standard position, which lies on the circle $x^2+y^2=r^2$ then $x=r\cos(\theta)$ and $y=r\sin(\theta)$. In the case $r>0$ theorem (1.11) is an immediate consequence of this observation along with the quotient identity $\tan(\theta)=\sin(\theta)/\cos(\theta)$.If $r<0$, then we know an alternate representation for $(r;\theta)$ is $(-r;\theta+\pi)$. Since $\cos(\theta+\pi)=-\cos(\theta)$ and $\sin(\theta+\pi)=-\sin(\theta)$, applying the theorem to $(-r;\theta+\pi)$ gives

$$x=(-r)\cos(\theta+\pi)=(-r)(-cos(\theta))=rcos(\theta)$$

and

$$y=(-r)\sin(\theta+\pi)=(-r)(-\sin(\theta))=r\sin(\theta).$$

Moreover, $x^2+y^2=(-r)^2=r^2$, and $y/x=\tan(\theta+\pi)=\tan(\theta)$, so the theorem is true in this case, too. The remaining case is $r=0$, in which case $(r;\theta)=(0;\theta)$ is the pole. Since the pole is identified with the origin $(0;0)$ in rectangular coordinates, the theorem in this case amounts to checking '$0=0$.'
■

The following example puts Theorem 1.11 to good use.

Example 1.12 *Convert each point in rectangular coordinates given below into polar coordinates with $r \geq 0$ and $0 \leq \theta < 2\pi$. Use exact values if possible and round any approximate values to two decimal places. Check your answer by converting them back to rectangular coordinates.*

1. $P(2, -2\sqrt{3})$ 2. $Q(-3, -3)$ 3. $R(0, -3)$ 4. $S(-3, 4)$.

Solution:

1. Even though we are not explicitly told to do so, we can avoid many common mistakes by taking the time to plot the points before we do any calculations. Plotting $P(2, -2\sqrt{3})$ shows that it lies in Quadrant IV. With $x = 2$ and $y = -2\sqrt{3}$, we get $r^2 = x^2+y^2 = (2)^2+\left(-2\sqrt{3}\right)^2 = 4+12 = 16$ so $r = \pm 4$. Since we are asked for $r \geq 0$, we choose $r = 4$. To find θ, we have that

$$\tan(\theta) = \frac{y}{x} = \frac{-2\sqrt{3}}{2} = -\sqrt{3}.$$

This tells us θ has a reference angle of $\frac{\pi}{3}$ and since P lies in Quadrant IV, we know θ is a Quadrant IV angle. We are asked to have $0 \leq \theta < 2\pi$, so we choose $\theta = \frac{5\pi}{3}$. Hence, our answer is $\left(4, \frac{5\pi}{3}\right)$ .To check, we convert $(r;\theta) = \left(4, \frac{5\pi}{3}\right)$ back to rectangular coordinates and we find

$$x = r\cos(\theta) = 4\sin\left(\frac{5\pi}{3}\right) = 4\left(\frac{-\sqrt{3}}{2}\right) = -2\sqrt{3}$$

as required.

2. The point $Q(-3, -3)$ lies in Quadrant III. Using $x = y = -3$, we get $r^2 = (-3)^2 + (-3)^2 = 18$ so $r = \pm\sqrt{18} = \pm 3\sqrt{2}$. Since we are asked for $r \geq 0$, we choose $r = 3\sqrt{2}$. We find $\tan(\theta)= \frac{-3}{-3} = 1$, which means θ has a reference angle of $\frac{\pi}{4}$. Since Q lies in Quadrant III, we choose $\theta = \frac{5\pi}{4}$, which satisfies the requirement that $0 \leq \theta < 2\pi$. Our final answer is $(r;\theta) = (3\sqrt{2}, \frac{5\pi}{4})$. To check, we find

$$x = r\cos(\theta) = 3\sqrt{2}\cos(\frac{5\pi}{4}) = 3\sqrt{2}\left(-\frac{\sqrt{2}}{2}\right) = -3$$

and

$$y = r\sin(\theta) = 3\sqrt{2}\sin(\frac{5\pi}{4}) = 3\sqrt{2}\left(-\frac{\sqrt{2}}{2}\right) = -3,$$

so we are done.

3. The point $R(0, -3)$ lies along the negative y-axis. While we could go through the usual computations to find the polar form of R, in this case we can find the polar coordinates of R using the definition. Since the pole is identified with the origin, we can easily tell the point R is 3 units from the pole, which means in the polar representation $(r; \theta)$ of R we know $r = \pm 3$. Since we require $r \geq 0$, we choose $r = 3$. Concerning θ, the angle $\theta = \frac{3\pi}{2}$ satisfies $0 \leq \theta < 2\pi$ with its terminal side along the negative y-axis, so our answer is $\left(3, \frac{3\pi}{2}\right)$. To check, we note

$$x = r\cos(\theta) = 3\cos(\frac{3\pi}{2}) = (3)(0) = 0$$

and

$$y = r\sin(\theta) = 3\sin(\frac{3\pi}{2}) = 3(-1) = -3.$$

4. The point[4] $S(-3, 4)$ lies in Quadrant II. With $x = -3$ and $y = 4$, we get $r^2 = (-3)^2 + (4)^2 = 25$ so $r = \pm 5$. As usual, we choose $r = 5 \geq 0$ and proceed to determine θ. We have

$$\tan(\theta) = \frac{y}{x} = \frac{4}{-3} = -\frac{4}{3},$$

and since this isn't the tangent of one the common angles, we resort to using the arctangent function. Since θ lies in Quadrant II and must satisfy $0 \leq \theta < 2\pi$, we choose $\theta = \pi - \arctan\left(\frac{4}{3}\right)$ radians. Hence, our answer is $(r; \theta) = \left(5, \pi - \arctan\left(\frac{4}{3}\right)\right) \simeq (5, 2.21)$. To check our answers requires a bit of tenacity since we need to simplify expressions of the form:

$$\cos\left(\pi - \arctan\left(\frac{4}{3}\right)\right) \quad \text{and} \quad \sin\left(\pi - \arctan\left(\frac{4}{3}\right)\right).$$

These are good review exercises and are hence left to the reader. We find

$$\cos\left(\pi - \arctan\left(\frac{4}{3}\right)\right) = -\frac{3}{5} \quad \text{and} \quad \sin\left(\pi - \arctan\left(\frac{4}{3}\right)\right) = \frac{4}{5}$$

so that

$$x = r\cos(\theta) = (5)(\frac{-3}{5}) = -3$$

and

$$y = r\sin(\theta) = (5)(\frac{4}{5}) = 4$$

which confirms our answer. □

---

[4] Skip this example if you are not familiar with the Inverse Trig Functions.

1.8 Exponential form

Let r and θ be polar coordinates of the point (x, y) that corresponds to a *nonzero* complex number $z = x + iy$. Since $x = r\cos\theta$ and $y = r\sin\theta$, the number z can be written in polar form as

$$z = r(\cos\theta + i\sin\theta). \tag{1.48}$$

If $z = 0$, the coordinate θ is undefined; and so it is understood that $z \neq 0$ whenever polar coordinates are used.

In complex analysis, the real number r is not allowed to be negative and is the length of the radius vector for z ; that is, $r = |z|$. The real number θ represents the angle, measured in radians, that z makes with the positive real axis when z is interpreted as a radius vector (Fig. 1.8). As in calculus, θ has an infinite number of possible values, including negative ones, that differ by integral multiples of 2π . Those values can be determined from the equation $\tan\theta = y/x$, where the quadrant containing the point corresponding to z must be specified. Each value of θ is called an *argument* of z, and the set of all such values is denoted by arg z. The *principal value* of arg z, denoted by Arg z, is that unique value θ such that $-\pi < \theta \leq \pi$. Evidently, then,

$$\arg z = \operatorname{Arg} z + 2n\pi \quad (n = 0, \pm 1, \pm 2, ...). \tag{1.49}$$

Also, when z is a negative real number, $\operatorname{Arg} z$ has value π , not $-\pi$.

Example 1.13 *The complex number $-1-i$, which lies in the third quadrant, has principal argument $-3\pi/4$. That is,*

$$3\pi \operatorname{Arg}(-1-i) = -\frac{3\pi}{4}.$$

It must be emphasized that because of the restriction $-\pi < \theta \leq \pi$ of the principal argument θ, it is not true that $\operatorname{Arg}(-1-i) = 5\pi/4$. According to equation (1.49), $\arg(-1-i) = -\frac{3\pi}{4} + 2n\pi \quad (n = 0, \pm 1, \pm 2, ...)$.

Note that the term $\operatorname{Arg} z$ on the right-hand side of equation (1.49) can be replaced by any particular value of $\arg z$ and that one can write, for instance,

$$\arg(-1-i) = \frac{5\pi}{4} + 2n\pi \quad (n = 0, \pm 1, \pm 2, ...).$$

The symbol $e^{i\theta}$, or $\exp(i\theta)$, is defined by means of *Euler's formula* as

$$e^{i\theta} = \cos\theta + i\sin\theta, \tag{1.50}$$

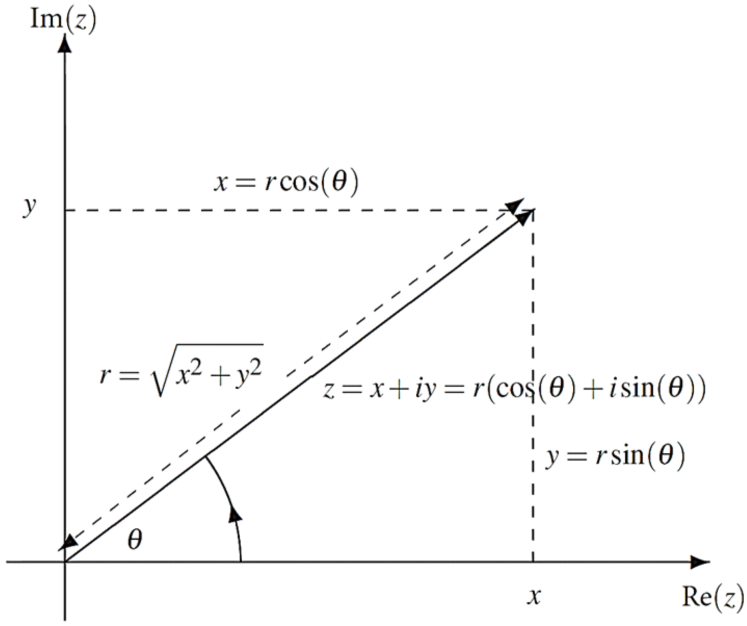


Fig. 1.8. Polar coordinates (r, θ) versus rectangular coordinates (x, y).

where θ is to be measured in radians. It enables one to write the polar form (1.48) more compactly in *exponential form* as

$$z = re^{i\theta} \tag{1.51}$$

The choice of the symbol $e^{i\theta}$ will be fully motivated later on.

Example 1.14 *The number* $-1-i$ *in Example 1.13 has exponential form*

$$-1 - i = \sqrt{2}\exp\left[i\left(-\frac{3\pi}{4}\right)\right]. \tag{1.52}$$

With the agreement that $e^{-i\theta} = e^{i(-\theta)}$, this can also be written $-1-i = \sqrt{2}e^{-i3\pi/4}$. Expression (1.52) is, of course, only one of an infinite number of possibilities for the exponential form of $-1-i$:

$$(n = 0, \pm 1, \pm 2, ...). \tag{1.53}$$

Note how expression (1.51) with $r = 1$ tells us that the numbers $e^{i\theta}$ lie on the circle centered at the origin with radius unity, as shown in Fig 1.9.

It is, for instance, geometrically obvious (without reference to Euler's formula) that

$$e^{i\pi} = -1, \qquad e^{-i\pi/2} = -i, \qquad \text{and} \qquad e^{-i4\pi} = 1.$$

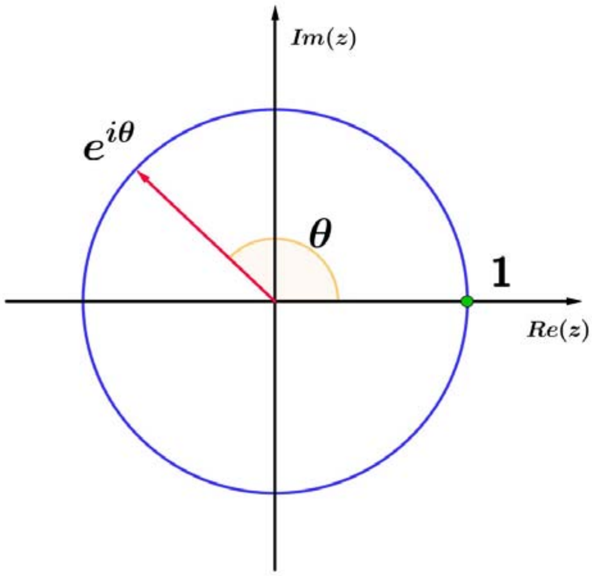


Fig. 1.9. The numbers $e^{i\theta}$ lie on the circle centered at the origin with radius unity.

Additionally, from Euler's formula we have

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

which together with (1.50) gives

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{and} \quad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}. \tag{1.54}$$

Note, too, that the equation

$$z = Re^{i\theta} \qquad (0 \leq \theta \leq 2\pi) \tag{1.55}$$

is a parametric representation of the circle $|z| = R$, centered at the origin with radius R. As the parameter θ increases from $\theta = 0$ to$\theta = 2\pi$, the point z starts from the positive real axis and traverses the circle once in the counterclockwise direction. More generally, the circle $|z - z_0| = R$, whose center is z_0 and whose radius is R, has the parametric representation

$$z = z_0 + Re^{i\theta} \qquad (0 \leq \theta \leq 2\pi). \tag{1.56}$$

This can be seen vectorially by noting that a point z traversing the circle $|z - z_0| = R$ once in the counterclockwise direction corresponds to the sum of the fixed vector z_0 and a vector of length R whose angle of inclination θ varies from $\theta = 0$ to $\theta = 2\pi$.

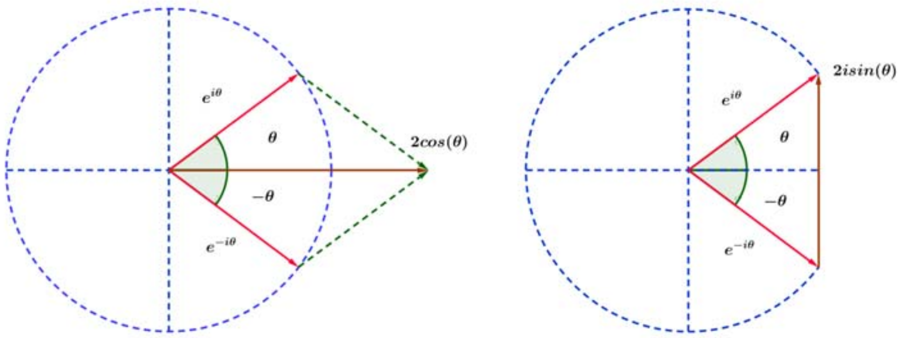


Fig. 1.10. Pictorial presentation of the equalities 1.54.

1.9 Products and powers in exponential form

Simple trigonometry tells us that $e^{i\theta}$ has the familiar additive property of the exponential function in calculus:

$$\begin{aligned} e^{i\theta_1}e^{i\theta_2} &= (\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) \\ &= (\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2) \\ &= \cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2) = e^{i(\theta_1+\theta_2)}. \end{aligned}$$

Thus, if $z_1 = r_1e^{i\theta_1}$ and $z_2 = r_2e^{i\theta_2}$, the product z_1z_2 has exponential form

$$z_1z_2 = r_1e^{i\theta_1}r_2e^{i\theta_2} = r_1r_2e^{i\theta_1}e^{i\theta_2} = (r_1r_2)e^{i(\theta_1+\theta_2)}. \qquad (1.57)$$

Furthermore,

$$\frac{z_1}{z_2} = \frac{r_1e^{i\theta_1}}{r_2e^{i\theta_2}} = \frac{r_1}{r_2}\frac{e^{i\theta_1}e^{-i\theta_2}}{e^{i\theta_2}e^{-i\theta_2}} = (\frac{r_1}{r_2})e^{i(\theta_1-\theta_2)} \qquad (z_2 \neq 0). \qquad (1.58)$$

Note how it follows from expression (1.58) that the inverse of any nonzero complex number $z = re^{i\theta}$ is

$$z^{-1} = \frac{1}{z} = \frac{1}{r}\frac{e^{i0}}{e^{i\theta}} = \frac{1}{r}e^{(0-\theta)} = \frac{1}{r}e^{-\theta}. \qquad (1.59)$$

Expressions (1.57), (1.58), and (1.59) are, of course, easily remembered by applying the usual algebraic rules for real numbers and e^x. Another important result that can be obtained formally by applying rules for real numbers to $z = re^{i\theta}$ is

$$z^n = r^ne^{in\theta} \quad (n = 0, \pm1, \pm2, ...). \qquad (1.60)$$

It is easily verified for positive values of n by mathematical induction. To be specific, we first note that it becomes $z = re^{i\theta}$ when $n = 1$. Next, we assume that it is valid when $n = m$, where m is any positive integer. In view of expression (1.57) for the product of two nonzero complex numbers in exponential form, it is then valid for $n = m + 1$:

$$z^{m+1} = z^m z = r^m e^{im\theta} re^{i\theta} = (r^m r)e^{i(m\theta+\theta)} = r^{m+1}e^{i(m+1)\theta}.$$

Expression (**??**) is thus verified when n is a positive integer. It also holds when $n = 0$, with the convention that $z^0 = 1$. If $n = -1, -2, ...$, on the other hand, we define z^n in terms of the multiplicative inverse of z by writing

$$z^n = (z^{-1})^m \quad \text{where} \quad m = -n = 1, 2, ...$$

Then, since equation (1.60) is valid for positive integers, it follows from the exponential form (1.59) of z^{-1} that

$$z^n = \left[\frac{1}{r}e^{i(-\theta)}\right]^m = \left(\frac{1}{r}\right)^m e^{im(-\theta)} = \left(\frac{1}{r}\right)^{-n} e^{i(-n)(-\theta)} = r^n e^{in\theta}$$

for $n = -1, -2, ...$ Expression (1.60) is now established for all integral powers. Expression (1.60) can be useful in finding powers of complex numbers even when they are given in rectangular form and the result is desired in that form.

In order to put $(\sqrt{3} + 1)^7$ in rectangular form, one need only write

$$(\sqrt{3} + i)^7 = (2e^{i\pi/6})^7 = 2^7 e^{i7\pi/6} = (2^6 e^{i\pi})(2e^{i\pi/6}) = -64(\sqrt{3} + i).$$

Finally, we observe that if $r = 1$, equation (1.60) becomes

$$(e^{i\theta})^n = e^{in\theta} \qquad (n = 0, \pm 1, \pm 2, ...). \tag{1.61}$$

When written in the form

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta \qquad (n = 0, \pm 1, \pm 2, ...), \tag{1.62}$$

this is known as *de Moivre's formula*. The following example uses a special case of it.

Example 1.15 *Formula (1.62) with $n = 2$ tells us that*

$$(\cos\theta + i\sin\theta)^2 = \cos 2\theta + i\sin 2\theta,$$

or

$$\cos^2\theta - \sin^2\theta + i2\sin\theta\cos\theta = \cos 2\theta + i\sin 2\theta.$$

By equating real parts and then imaginary parts here, we have the familiar trigonometric identities

$$\cos 2\theta = \cos^2\theta - \sin^2\theta, \qquad \sin 2\theta = 2\sin\theta\cos\theta.$$

1.10 Arguments of products and quotiens

If $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$, the expression (1.57)

$$z_1 z_2 = (r_1 r_2) e^{i(\theta_1+\theta_2)} \tag{1.63}$$

in Sec. 1.9 can be used to obtain an important identity involving arguments:

$$\arg(z_1 z_2) = \arg z_1 + \arg z_2. \tag{1.64}$$

This result is to be interpreted as saying that if values of two of the three (multiple-valued) arguments are specified, then there is a value of the third such that the equation holds.

We start the verification of statement (1.64) by letting θ_1 and θ_2 denote any values of arg z_1 and arg z_2, respectively. Expression (1.63) then tells us that $\theta_1 + \theta_2$ is a value of $\arg(z_1 z_2)$. (See Fig. 1.11.) If, on the other hand, values of $\arg(z_1 z_2)$ and arg z_1 are specified, those values correspond to particular choices of n and n_1 in the expressions

$$\arg(z_1 z_2) = (\theta_1 + \theta_2) + 2n\pi \qquad (n = 0, \pm 1, \pm 2, ...)$$

and

$$\arg z_1 = \theta_1 + 2n_1\pi \qquad (n_1 = 0, \pm 1, \pm 2, ...).$$

Since

$$(\theta_1 + \theta_2) + 2n\pi = (\theta_1 + 2n_1\pi) + [\theta_2 + 2(n - n_1)\pi],$$

equation (1.64) is evidently satisfied when the value

$$\arg z_2 = \theta_2 + 2(n - n_1)\pi$$

is chosen. Verification when values of $\arg(z_1 z_2)$ and arg z_2 are specified follows by symmetry.

Statement (1.64) is sometimes valid when arg is replaced everywhere by Arg . But, as the following example illustrates, that is *not always* the case.

Example 1.16 *When* $z_1 = -1$ *and* $z_2 = i$,

$$\mathrm{Arg}(z_1 z_2) = \mathrm{Arg}(-i) = -\frac{\pi}{2} \qquad \textit{but} \qquad \mathrm{Arg}\, z_1 + \mathrm{Arg}\, z_2 = \pi + \frac{\pi}{2} = \frac{3\pi}{2}.$$

If, however, we take the values of $\arg z_1$ *and* $\arg z_2$ *just used and select the value*

$$\mathrm{Arg}(z_1 z_2) + 2\pi = -\frac{\pi}{2} + 2\pi = \frac{3\pi}{2}$$

of $\arg(z_1 z_2)$*, we find that equation (1.64) is satisfied.* □

Statement (1.64) tells us that

$$\arg(\frac{z_1}{z_2}) = \arg(z_1 z_2^{-1}) = \arg(z_1) + \arg(z_2^{-1});$$

and, since (Sec. 1.9)

$$z_2^{-1} = \frac{1}{r_2} e^{-i\theta_2}$$

one can see that

$$\arg(z_2^{-1}) = -\arg(z_2). \tag{1.65}$$

Hence (see Fig. 1.12.)

$$\arg(\frac{z_1}{z_2}) = \arg(z_1) - \arg(z_2). \tag{1.66}$$

Statement (1.65) is, of course, to be interpreted as saying that the set of all values on the left-hand side is the same as the set of all values on the right-hand side. Statement (1.66) is, then, to be interpreted in the same way that statement (1.64) is.

Example 1.17 *In order to find the principal argument* $\mathrm{Arg}\, z$ *when*

$$z = \frac{-2}{1+\sqrt{3}i}$$

observe that

$$\arg z = \arg(-2) - \arg(1+\sqrt{3}i).$$

Since

$$\mathrm{Arg}(-2) = \pi \quad \textit{and} \quad \mathrm{Arg}(1+\sqrt{3}i) = \frac{\pi}{3}$$

one value of $\arg z$ *is* $2\pi/3$*; and, because* $2\pi/3$ *is between* $-\pi$ *and* π *, we find that* $\mathrm{Arg}\, z = 2\pi/3$*. When* z_1 *and* z_2 *are multiplied* $|z_1 z_2| = r_1 r_2$ *and* $\arg(z_1 z_2) = \theta_1 + \theta_2$.

Example 1.18 *Use polar forms of the complex numbers*

$$z_1 = 1 + \sqrt{3}i \quad \textit{and} \quad z_2 = \sqrt{3} + i$$

to compute $z_1 z_2 =$ *and* z_1/z_2.

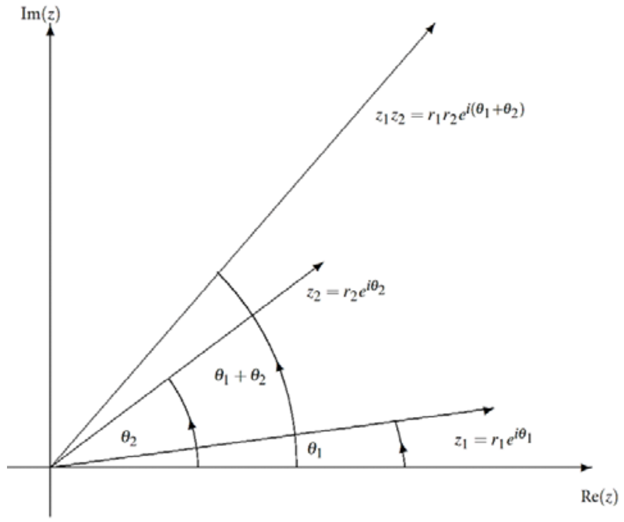


Fig. 1.11. Visualization of products. When z_1 and z_2 are multiplied $|z_1 z_2| = r_1 r_2$ and $\arg(z_1 z_2) = \theta_1 + \theta_2$.

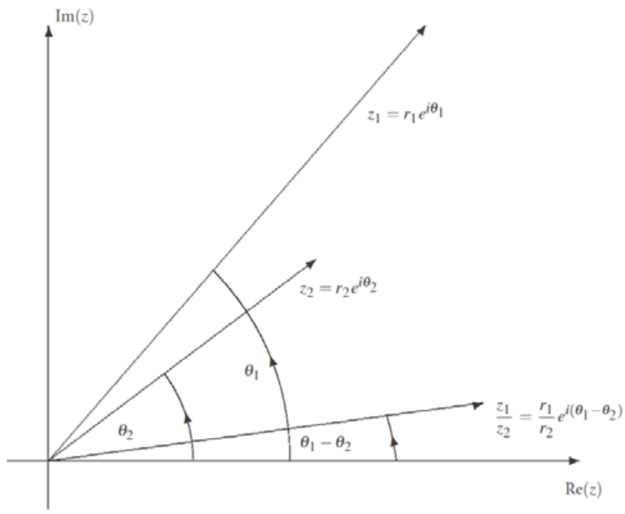


Fig. 1.12. Visualization of quotients. We divide lengths and subtract angles for the quotines of complex numbers z_1 and z_2, so $\left|\frac{z_1}{z_2}\right| = \frac{r_1}{r_2}$ and $\arg(\frac{z_1}{z_2}) = \theta_1 - \theta_2$.

Solution: Polar forms of these complex numbers are

$$z_1 = 2\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right) \quad \text{and} \quad z_2 = 2\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right)$$

(verify). Thus, it follows from (1.63) that

$$\begin{aligned} z_1 z_2 &= 4\left[\cos\left(\frac{\pi}{3}+\frac{\pi}{6}\right) + i\sin\left(\frac{\pi}{3}+\frac{\pi}{6}\right)\right] \\ &= 4\left[\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right] = 4i. \end{aligned}$$

and from (1.66)

$$\begin{aligned} \frac{z_1}{z_2} &= 1\left[\cos\left(\frac{\pi}{3}-\frac{\pi}{6}\right) + i\sin\left(\frac{\pi}{3}-\frac{\pi}{6}\right)\right] \\ &= \cos\left(\frac{\pi}{6}\right) + i\sin\left(\frac{\pi}{6}\right) = \frac{\sqrt{3}}{2} + \frac{1}{2}i. \end{aligned}$$

As a check we can calculate $z_1 z_2 =$ and z_1/z_2 directly

$$z_1 z_2 = \left(\sqrt{3}+i\right)\left(i\sqrt{3}+1\right) = \sqrt{3} + i + 3i + \sqrt{3}i^2 = 4i,$$

$$\frac{z_1}{z_2} = \frac{(1+i\sqrt{3})}{\sqrt{3}+i} = \frac{(1+i\sqrt{3})\left(\sqrt{3}-i\right)}{\left(\sqrt{3}+i\right)\left(\sqrt{3}-i\right)} = \frac{\sqrt{3} - i + 3i - \sqrt{3}i^2}{3 - i^2} = \frac{\sqrt{3}}{2} + \frac{1}{2}i$$

which agrees with the results obtained using polar forms. □

1.11 Roots of complex numbers

Consider now a point $z = re^{i\theta}$, lying on a circle centered at the origin with radius r (see Fig. 1.13).

As θ is increased, z moves around the circle in the counterclockwise direction. In particular, when θ is increased by 2π, we arrive at the original point; and the same is true when θ is decreased by 2π . It is, therefore, evident from (Fig. 1.13) that two nonzero complex numbers

$$z_1 = r_1 e^{i\theta_1} \qquad \text{and} \qquad z_2 = r_2 e^{i\theta_2}$$

are equal if and only if

$$r_1 = r_2 \qquad \text{and} \qquad \theta_1 = \theta_2 + 2k\pi,$$

where k is some integer ($k = 0, \pm 1, \pm 2, \ldots$).

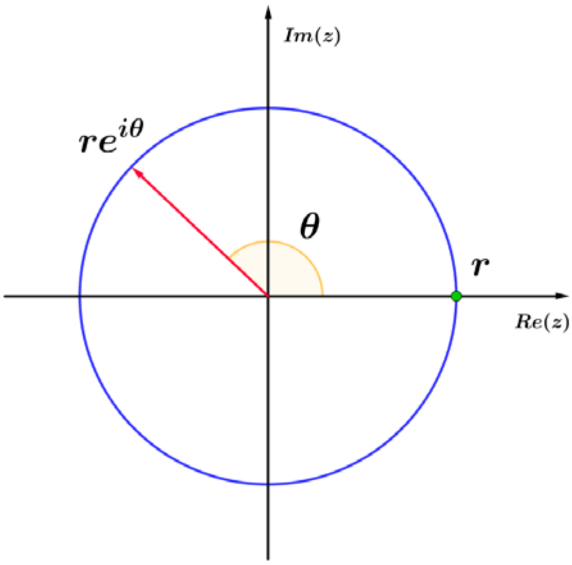


Fig. 1.13. A point $z = re^{i\theta}$, lying on a circle centered at the origin with radius r.

This observation, together with the expression $z^n = r^n e^{in\theta}$ in Sec. 1.9 for integral powers of complex numbers $z = re^{i\theta}$, is useful in finding the n-th roots of any nonzero complex number $z_0 = r_0 e^{i\theta_0}$, where n has one of the values $n = 2, 3, ...$ The method starts with the fact that an $n-$th root of z_0 is a nonzero number $z = re^{i\theta}$ such that $z^n = z_0$, or

$$r^n e^{in\theta} = r_0 e^{i\theta_0}.$$

According to the statement in italics just above, then,

$$r^n = r_0 \qquad \text{and} \qquad n\theta = \theta_0 + 2k\pi,$$

where k is any integer($k = 0, \pm 1, \pm 2, ...$). So $r = \sqrt[n]{r_0}$, where this radical denotes the unique positive $n-$th root of the positive real number r_0, and

$$\theta = \frac{\theta_0 + 2k\pi}{n} = \frac{\theta_0}{n} + \frac{2k\pi}{n} \qquad (k = 0, \pm 1, \pm 2, ...).$$

Consequently, the complex numbers

$$z = \sqrt[n]{r_0} \exp\left[i\left(\frac{\theta_0}{n} + \frac{2k\pi}{n}\right)\right] \qquad (k = 0, \pm 1, \pm 2, ...)$$

are the $n-$th roots of z_0. We are able to see immediately from this exponential form of the roots that they all lie on the circle $|z| = \sqrt{n}\, r_0$ about the origin and

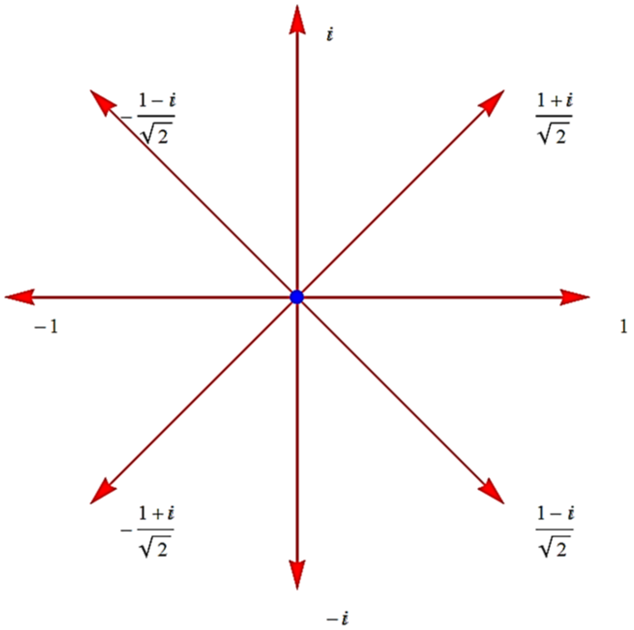


Fig. 1.14. Roots of unity ($n = 8$).

are equally spaced every $2\pi/n$ radians, starting with argument θ_0/n. Evidently, then, all of the distinct roots are obtained when $k = 0, 1, 2, ..., n-1$, and no further roots arise with other values of k. We let c_k ($k = 0, 1, 2, ..., n-1$) denote these *distinct* roots**??** and write

$$c_k = \sqrt[n]{r_0} \exp\left[i\left(\frac{\theta_0}{n} + \frac{2k\pi}{n}\right)\right] \qquad (k = 0, 1, 2, ..., n-1). \tag{1.67}$$

The number $\sqrt[n]{r_0}$ is the length of each of the radius vectors representing the n roots. The first root c_0 has argument θ_0/n; and the two roots when $n = 2$ lie at the opposite ends of a diameter of the circle $\{z : |z| = \sqrt[n]{r_0}\}$, the second root being $-c_0$. When $n \geq 3$, the roots lie at the vertices of a regular polygon of n sides inscribed in that circle (see Fig. **??**). We shall let $z_0^{1/n}$ denote the *set* of n-th roots of z_0. If, in particular, z_0 is a positive real number r_0, the symbol $r_0^{1/n}$ denotes the entire set of roots; and the symbol $\sqrt[n]{r_0}$ in expression (1.67) is reserved for the one positive root. When the value of θ_0 that is used in expression (**??**) is the principal value of arg z_0 ($-\pi < \theta_0 \leq \pi$), the number c_0 is referred to as the *principal root*. Thus when z_0 is a positive real number r_0, its principal root is $\sqrt[n]{r_0}$.

Observe that if we write expression (**??**) for the roots of z_0 as

$$c_k = \sqrt[n]{r_0}\exp\left[i\left(\frac{\theta_0}{n} + \frac{2k\pi}{n}\right)\right] \qquad (k = 0, 1, 2, ..., n-1)$$

and also write

$$\omega_n = \exp\left(i\frac{2\pi}{n}\right) \tag{1.68}$$

it follows from property (**??**), Sec. 1.9.) of $e^{i\theta}$ that

$$\omega_n^k = \exp\left(i\frac{2\pi k}{n}\right) \qquad (k = 0, 1, 2, ..., n-1) \tag{1.69}$$

and hence that

$$c_k = c_0\omega_n^k \qquad (k = 0, 1, 2, ..., n-1) \tag{1.70}$$

The number c_0 here can, of course, be replaced by any particular n-th root of z_0, since ω_n represents a counterclockwise rotation through $2\pi/n$ radians. Finally, a convenient way to remember expression (1.67) is to write z_0 in its most general exponential form

$$z_0 = r_0 e^{i(\theta_0 + 2k\pi)} \qquad (k = 0, \pm 1, \pm 2, ...)$$

and to *formally* apply laws of fractional exponents involving real numbers, keeping in mind that there are precisely n roots:

$$z_0^{1/n} = \left[r_0 e^{i(\theta_0 + 2k\pi)}\right]^{1/n} = \sqrt[n]{r_0}\exp\left[i\left(\frac{\theta_0}{n} + \frac{2k\pi}{n}\right)\right] \quad (k = 0, 1, 2, ..., n-1).$$

The examples in the next section serve to illustrate this method for finding roots of complex numbers.

For a positive integer n, the complex numbers ω_n^k from (1.69) i.e.

$$\{1, \omega_n, \omega_n^2, ..., \omega_n^{n-1}\}$$

are called the n-th *roots of unity* because they represent all solutions to $z^n = 1$. Geometrically, they are the vertices of a regular polygon of n sides as depicted in Figure (**??**) for $n = 12$.The roots of unity are cyclic in the sense that if $k \geq n$, then $\omega^k = \omega^{k(\bmod n)}$, where $k(\bmod n)$ denotes the remainder when k is divided by n–for example, when $n = 6$, $\omega_6^6 = 1$, $\omega_6^7 = \omega_6$, $\omega_6^8 = \omega_6^2$, $\omega_6^9 = \omega_6^3, ...$

Example 1.19 *The three cube roots of* $1 = \cos 0 + i \sin 0$ *are*

$$\begin{aligned} \cos\frac{0}{3} + i\sin\frac{0}{3} &= 1, \\ \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} &= -\frac{1}{2} + \frac{1}{2}i\sqrt{3}, \\ \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} &= -\frac{1}{2} - \frac{1}{2}i\sqrt{3}, \end{aligned}$$

which correspond to taking $k = 0, 1$ *and* 2 *in (1.69).*□

Example 1.20 *We solve the equation* $z^6 = -1$ *to find the* 6*th roots of* -1*. Writing* $z = re^{i\theta}$*, we have*

$$z^6 = (re^{i\theta})^6 = r^6 e^{i6\theta},$$

and

$$-1 = e^{i\pi} = e^{i(\pi+2k\pi)} \quad \textit{for} \quad k \in \mathbb{Z}.$$

So we need to solve

$$r^6 e^{i6\theta} = e^{i(\pi+2k\pi)}.$$

Using the fact that r *is a real positive number, we have* $r = 1$ *and* $6\theta = \pi + 2k\pi$*, so*

$$\theta = \frac{\pi}{6} + \frac{2\pi k}{6}.$$

This will give the six distinct complex roots by taking $k = 0, 1, 2, 3, 4, 5$. □

Example 1.21 *Show that all roots of* $(z+1)^5 + z^5 = 0$ *lie on line* $x = -\frac{1}{2}$.

Solution: Each root of this equation is of course different from zero. Dividing both sides of the equation by z^5 we have an equivalent relation

$$1 + \left(1 + \frac{1}{z}\right)^5 = 0.$$

Therefore

$$1 + \frac{1}{z} = e^{i(\pi+2k\pi)/5}, \qquad k = 0, ...4;$$

and thus

$$z^{-1} = -1 + e^{i(\pi+2k\pi)/5}.$$

Designating $\theta = (\pi + 2k\pi)/5$ we can write (for each of the allowed value of k)

$$\begin{aligned} z^{-1} &= \cos\theta + i\sin\theta - 1 \\ &= -2\sin^2\frac{\theta}{2} + 2i\sin\frac{\theta}{2}\cos\frac{\theta}{2} \\ &= -2\sin\frac{\theta}{2}\left[\sin\frac{\theta}{2} - i\cos\frac{\theta}{2}\right] \\ &= -2\sin\frac{\theta}{2}\left[\cos\left(\frac{\pi}{2} - \frac{\theta}{2}\right) - i\sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right)\right]. \end{aligned}$$

Thus

$$z^{-1} = -2\sin\frac{\theta}{2}e^{-i\left(\frac{\pi}{2} - \frac{\theta}{2}\right)}$$

and

$$z = \frac{1}{-2\sin\frac{\theta}{2}}\left[\cos\left(\frac{\pi}{2} - \frac{\theta}{2}\right) + i\sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right)\right].$$

Hence

$$\mathrm{Re}(z) = -\frac{1}{-2\sin\frac{\theta}{2}}\sin\frac{\theta}{2} = -\frac{1}{2} \quad \text{for} \quad k = 0, ...4$$

NOTE: It is not good enough to just say

$$(1+z)^5 = -z^5$$

hence

$$1 + z = -z, \quad 2z = -1, \quad z = -\frac{1}{2}.$$

This only shows that $z = (-1/2, 0)$ is one root.

□

1.12 Summing trigonometric series

Below are two different series. Which is it easier to sum?

$$C_1 = 1 + \cos\theta + \cos 2\theta + \cos 3\theta + ... + \cos n\theta, \tag{1.71}$$

$$C_2 = 1 + \cos\theta + \cos^2\theta + \cos^3\theta + ... + \cos^n\theta. \tag{1.72}$$

The answer is that C_2 is simply a geometric series, whose sum can be written down immediately as

$$C_2 = \frac{1 - \cos^{n+1}\theta}{1 - \cos\theta}. \tag{1.73}$$

The first looks more complicated. However, we remember that de Moivre's formula (1.62, p. 34) gives us a recipe for converting the sines and cosinesof multiple angles (like 3θ) to the corresponding powers. Thus the problem of summing C_1 would have been easier if it had been coupled with that of summing

$$S_1 = \sin\theta + \sin 2\theta + \sin 3\theta + ... + \sin n\theta \tag{1.74}$$

since in that case we can multiply S_1 by i and add it to C_1 to obtain

$$\begin{aligned} & C_1 + iS_1 \\ = \; & 1 + (cos\theta + i\sin\theta) + (\cos 2\theta + i\sin 2\theta) + ... + (\cos n\theta + i\sin n\theta) \\ = \; & 1 + z + z_2 + ... + z^n \end{aligned}$$

where $z = \cos\theta + i\sin\theta \neq 1$, and we have made use of de Moivre's formula. Thus we have obtained $C_1 + iS_1$ as a geometric series which we can sum to give

$$\begin{aligned} & C_1 + iS_1 \\ = \; & \frac{1 - z^{n+1}}{1 - z} \\ = \; & \frac{(1 - e^{i(n+1)\theta})}{(1 - e^{i\theta})} \\ = \; & \frac{e^{i(n+1)\theta/2}\sin[(n+1)\theta/2]}{e^{i\theta/2}\sin(\theta/2)} \end{aligned} \tag{1.75}$$

where we have used de Moivre's formula along with formulas (1.54) to conclude, that for an arbitraty ϕ

$$\begin{aligned} 1 - e^{i\phi} &= e^{i\phi/2}e^{-i\phi/2} - e^{i\phi/2}e^{i\phi/2} \\ &= 2ie^{i\phi/2}\frac{\left(e^{i\phi/2} - e^{-i\phi/2}\right)}{2i} \\ &= 2ie^{i\phi/2}\sin(\theta/2). \end{aligned}$$

Next this relationship was applied to the numerator and denominator in (1.75), together with the natural simplification ($2i$ cancels).

Finally we extract C_1 and S_1 by comparing the real and the imaginary parts of equation (1.75) and obtain

$$C_1 = \cos(n\theta/2)\sin[(n+1)\theta/2]/\sin(\theta/2) \qquad (\theta \neq 2k\pi) \tag{1.76}$$

and

$$S_1 = \sin(n\theta/2)\sin[(n+1)\theta/2]/\sin(\theta/2) \qquad (\theta \neq 2k\pi). \tag{1.77}$$

Thus what is needed is the imagination to see that, if we are presented with the series C_1, which involves only real terms, we can make it easier by inventing the series S_1 and turning the problem into one containing complex numbers. In our series for C_1 above, the coefficients of the various terms $\cos k\theta$ were all unity. If the coefficients instead are some simple function of k, we may still be able to add up the series for z, and hence solve the problem. An example of this type appears in Problem (**??**)

1.13 Roots of polynomials

As it was mentioned earlier *The Fundamental Theorem of Algebra* asserts that a polynomial of degree n with complex coefficients has n complex roots (not necessarily distinct), and can therefore be factorised into n linear factors. If the coefficients are restricted to real numbers, the polynomial can be factorised into a product of linear and irreducible quadratic factors over $\mathbb{R}$ and into a product of linear factors over $\mathbb{C}$. However, we note the following useful result:

Theorem 1.22 *Complex roots of polynomials with real coefficients appear in conjugate pairs.*

Proof. Let $P(x) = a_0 + a_1x + \cdots + a_nx^n$, $a_i \in \mathbb{R}$, be a polynomial of degree n. We shall show that if z is a root of $P(x)$, then so is $\bar{z}$. ■

Let z be a complex number such that $P(z) = 0$, then

$$a_0 + a_1z + a_2z^2 \cdots + a_nz^n = 0$$

Conjugating both sides of this equation, we get

$$\overline{a_0 + a_1z + a_2z^2 \cdots + a_nz^n} = \bar{0}$$

Since 0 is a real number, it is equal to its complex conjugate. We now use the following properties of the complex conjugate: that the complex conjugate of the sum is the sum of the conjugates, and the complex conjugate of a product is the product of the conjugates.We have

$$\bar{a}_0 + \overline{a_1z} + \overline{a_2z^2} \cdots + \overline{a_nz^n} = 0$$

and

$$\bar{a}_0 + \bar{a}_1\bar{z} + \bar{a}_2\bar{z}^2 \cdots + \bar{a}_n\bar{z}^n = 0$$

Since the coefficients ai are real numbers, this becomes

$$a_0 + a_1\bar{z} + a_2\bar{z}^2 \cdots + a_n\bar{z}^n = 0$$

That is, $P(\bar{z}) = 0$, so the number z is also a root of $P(x)$.

The fact that every complex number has a square root is easily seen from the polar form: $\sqrt{r}e^{i(\theta/2)}$ is a square root of $re^{i\theta}$. From this we may deduce that every quadratic equation

$$az^2 + bz + c = 0,$$

where $a, b, c \in \mathbb{C}$ and $a \neq 0$ has a solution in $\mathbb{C}$. The procedure, by "completing the square", and the resulting formula

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

are just the same as for real quadratic equations.

Example 1.23 *Find the roots of the equation*

$$z^2 + 2iz + (2 - 4i) = 0.$$

Solution: By the standard formula, the solution of the equation is

$$\begin{aligned} & \frac{1}{2}\left(-2i \pm \sqrt{(-2i)^2 - 4(2-4i)}\right) \\ = & \frac{1}{2}\left(-2i \pm \sqrt{(-12+16i)}\right) \\ = & -i \pm \sqrt{-3+4i}. \end{aligned}$$

Observe now that $(1+2i)^2 = -3+4i$, and so the solution is $z = -i \pm (1+2i) = 1+i \quad \text{or} \quad -1-3i$ Note that this time the roots do not appear in conjugate pair as the coefficients of our polynomial were not real. □

Example 1.24 *Let us consider the polynomial*

$$x^3 - 2x^2 - 2x - 3 = (x-3)(x^2 + x + 1).$$

If

$$w = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$$

then

$$x^3 - 2x^2 - 2x - 3 = (x-3)(x-w)(x-\bar{w}).$$

□

Example 1.25 *Find all the roots of the equation* $z^4 + 1 = 0$*. Factorize the polynomial in* $\mathbb{C}$*, and also in* $\mathbb{R}$*.*

Solution: $z^4 = -1 = e^{i\pi r}$ if and only if $z = e^{\pm \pi i/4}$ or $e^{\pm 3\pi i/4}$. The roots all lie on the unit circle, and are equally spaced . In $\mathbb{C}$ the factorizations is

$$= (z - e^{\pi i/4})(z - e^{-\pi i/4})(z - e^{3\pi i/4})(z - e^{-3\pi i/4}).$$

Combining conjugate factors, we obtain the factorizations in $\mathbb{R}$:

$$\begin{aligned} z^4 + 1 &= (z^2 - 2z\cos(\pi/4) + 1)(z^2 - 2z\cos(3\pi/4) + 1) \\ &= (z^2 - \sqrt{2}z + 1)(z^2 + \sqrt{2}z + 1). \end{aligned}$$

□

$$\omega_k := e^{i2k\pi/n} \qquad k = (0, 1, ..., n-1).$$

1.14 Complex Numbers and Geometry

Several features of complex numbers make them extremely useful in plane geometry. For example, probably one of the more popular "math facts" that the central angle in a circle is twice the inscribed angle subtended by the same arc can be easily proved with complex numbers. The original formulation comes from Euclid:

> *In a circle the angle at the center is double the angle at the circumference when the angles have the same circumference as base.*

An important corollary of this fact is that, in a circle, all inscribed angles subtended by the same arc are equal. The proof is based on the following Fig. (1.15).

Without loss of generality, the circle is assumed to have radius 1 and be centered at the origin. The points are *identified* with complex numbers, so that, say, the equation of the circle is $z = e^{i\psi}$ where ψ is a real number, the angle measured from the horizontal x-axis. Let B correspond to $\psi = 0$, A to $\psi = \phi$, and P to $\psi = \vartheta$. Thus, $B = 1$, $A = e^{i\phi}$, $P = e^{i\vartheta}$. We find, for vectors,

$$\begin{aligned} AP &= e^{i\vartheta} - e^{i\phi}, \\ BP &= e^{i\vartheta} - 1. \end{aligned}$$

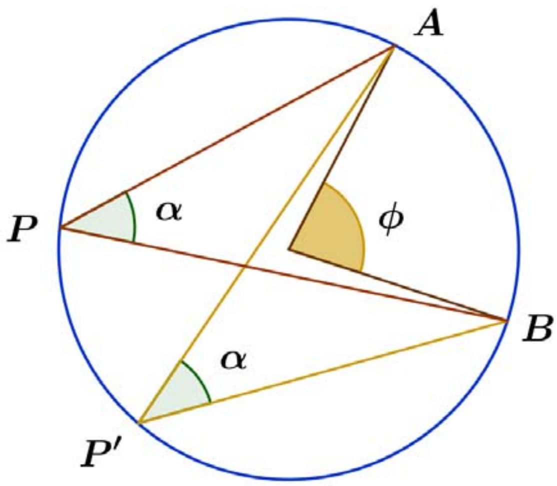


Fig. 1.15. Central angle in a circle is twice the inscribed angle subtended by the same arc.

The argument of the ratio of this two complex numbers is $\angle APB = \alpha$. Dividing the ratio by its conjugate eliminates the unimportant length of the ratio, but doubles the argument. Thus we have

$$
\begin{aligned}
e^{2i\alpha} &= \frac{e^{i\vartheta} - e^{i\phi}}{e^{i\vartheta} - 1} : \frac{e^{-i\vartheta} - e^{-i\phi}}{e^{-i\vartheta} - 1} \\
&= \frac{e^{i\vartheta} - e^{i\phi}}{e^{i\vartheta} - 1} \cdot \frac{e^{-i\vartheta} - 1}{e^{-i\vartheta} - e^{-i\phi}} \\
&= \frac{e^{i\vartheta} - e^{i\phi}}{e^{i\vartheta} - 1} \cdot \frac{e^{i\phi}\left(1 - e^{-i\vartheta}\right)}{e^{i\phi} = e^{i\vartheta}} \\
&= e^{i\phi}
\end{aligned}
$$

Assuming all angles are between 0 and π, $2\alpha = \phi$.

1.15 Fractals

With computers, we can generate beautiful art from complex numbers. These designs, some of which you can see on this page, are called *fractals*. Fractals are produced using an iteration process. This is where we start with a number and then feed it into a formula. We get a result and feed this result back into the formula, getting another result. And so on and so on...Fractals start with a

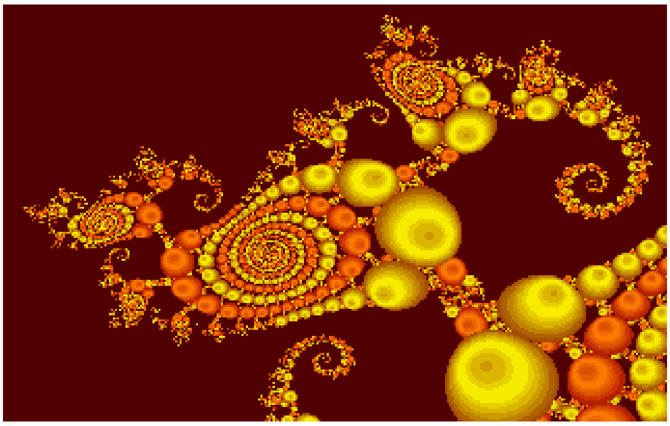


Fig. 1.16. Exemplary fractal.

complex number. Each complex number produced gives a value for each pixel on the screen. The higher the number of iterations, the better the quality of the image.

Common fractals are based on the *Julia Set* and the Mandelbrot Set. The Julia Set equation is:

$$Z_{n+1} = (Z_n)^2 + c.$$

For the Julia Set, the value of c remains constant and the value of Z_n changes. If we start with the complex number

$$Z_1 = 0.5 + 0.6i,$$

and let $c = 0.3$ and then feed this into the formula above, we have:

$$Z_2 = (0.5 + 0.6i)^2 + 0.3 = 0.19 + 0.6i$$

We now take this new answer and feed it back in:

$$Z_3 = (0.19 + 0.6i)^2 + 0.3 = -0.0239 + 0.228i$$

Continuing, we find that

$$Z_4 = 0.24858721 - 0.0108984i,$$

and

$$Z_5 = 0.3616768258 - 0.005418405698i$$

The Mandelbrot Set (discovered accidentally by an IBM computer programmer of that name) is the same as the Julia Set, but the value of c is allowed to change.

Fig. 1.17. Fractal fern.

Another famous example is the "fractal fern" (see Fig. 1.17). This is not a digital photograph - it is completely computer-generated by fractals.

Much solid and fascinating mathematics is involved in a proper study of fractals, but this is well beyond the scope of an introductory book[5].

Is There a Use for Any of This? Yes! For example US company called Fractal Antenna Systems, Inc. makes antenna arrays that use fractal shapes to get superior performance characteristics, because they can be packed so close together. More details can be found at: http://www.spacedaily.com/news/antenna-02d.html

[5] For a mathematical account of fractal sets, including Julia sets and the Mandelbrot set, see Kenneth Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, 2nd Edition, Wiley, 2003.

2

Systems of Linear Equations

A **linear equation:**

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

Example 2.1

- ***Linear:***

$$\begin{array}{ccc} 4x_1 - 5x_2 + 2 = x_1 & \textit{and} & x_2 = 2(\sqrt{6} - x_1) + x_3 \\ \downarrow & & \downarrow \\ \textit{rearranged} & & \textit{rearranged} \\ \downarrow & & \downarrow \\ 3x_1 - 5x_2 = -2 & & 2x_1 + x_2 - x_3 = 2\sqrt{6} \end{array}$$

- ***Not linear:***

$$4x_1 - 6x_2 = x_1x_2 \qquad \textit{and} \qquad x_2 = 2\sqrt{x_1} - 7$$

A **system of linear equations** (or a **linear system):**

A collection of one or more linear equations involving the same set of variables, say, $x_1, x_2, ..., x_n$.

A **solution** of a linear system:

A list $(s_1, s_2, ..., s_n)$ of numbers that makes each equation in the system true when the values $s_1, s_2, ..., s_n$ are substituted for $x_1, x_2, ..., x_n$, respectively.

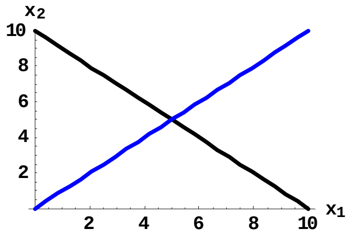
The **solution set:**
The set of all possible solutions of a linear system.

LINES, PLANES, HYPERPLANES.
The set of points in the plane satisfying $ax + by = c$ form a **line**.
The set of points in space satisfying $ax + by + cd = d$ form a **plane**.
The set of points satisfying $a_1 x_1 + ... + a_n x_n = a_0$ define a set called a **hyperplane** in n-dimensional space.

Example 2.2 *Two equations in two variables*

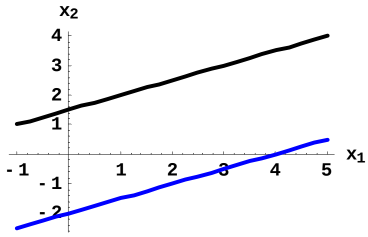
$$\begin{array}{rcrcr} x_1 & + & x_2 & = & 10 \\ -x_1 & + & x_2 & = & 0 \end{array}$$

:



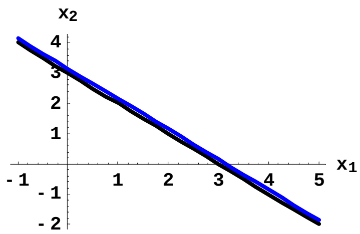
one unique solution

$$\begin{array}{rcrcr} x_1 & - & 2x_2 & = & -3 \\ 2x_1 & - & 4x_2 & = & 8 \end{array}$$



no solution

$$\begin{array}{rcrcr} x_1 & + & x_2 & = & 3 \\ -2x_1 & - & 2x_2 & = & -6 \end{array}$$



infinitely many solutions

SOLVE BY COMPUTER.
Use the computer. In Mathematica:

Solve[$\{5x - 2y + 7z == 15,\ 3x + 8y - z == -4,\ -9x + 6y + 10z == 7\}$, $\{x, y, z\}$]

But what did Mathematica do to solve this equation? We will look in this course at some efficient algorithms.

GEOMETRIC SOLUTION.
If a linear equation has only three variables, then they are usually denoted by x, y, and z with a linear equation written as

$$ax + by + cz = d$$

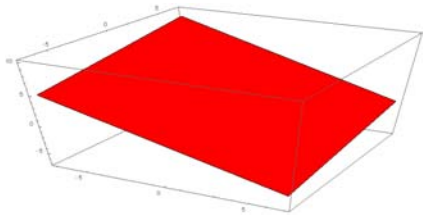


Fig. 2.1. Solution set for three variable linear equation.

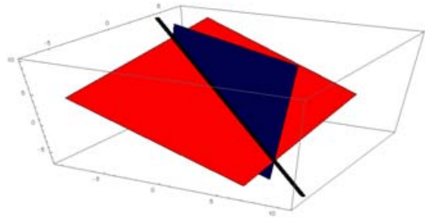


Fig. 2.2. The intersection of two planes is here a line.

for real constant a through d. The plot of the solution set for such a three-variable linear equation is a plane in space with coordinates x, y, and z. As an example, consider the following system of linear equations

$$\begin{aligned} 5x - 2y + 7z &= 15 \\ 3x + 8y - z &= -4 \\ -9x + 6y + 10z &= 7. \end{aligned}$$

Let's the plot of the solution set for the first three variable linear equation from that system (see Figure 2.1).

The second plane is the solution set to the second equation. To satisfy the first two equations means to be on the intersection of these two planes which is here a line (Figure 2.2).

To satisfy all three equations (in this case), we have to intersect the line with the plane representing the third equation which is a point (Figure 2.3).

Two equations could contradict each other. Geometrically this means that the two planes do not intersect. This is possible if they are parallel. Even without two planes being parallel, it is possible that there is no intersection between all three of them (Figure 2.4). Also possible is that we don't have enough equations (for example because two equations are the same) and that there are many solutions. Furthermore, we can have too many equations and the four planes would not intersect.

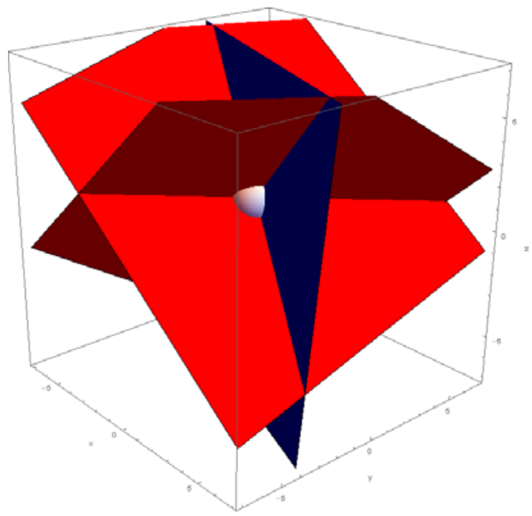


Fig. 2.3. Unique solution for the system of three linear equations

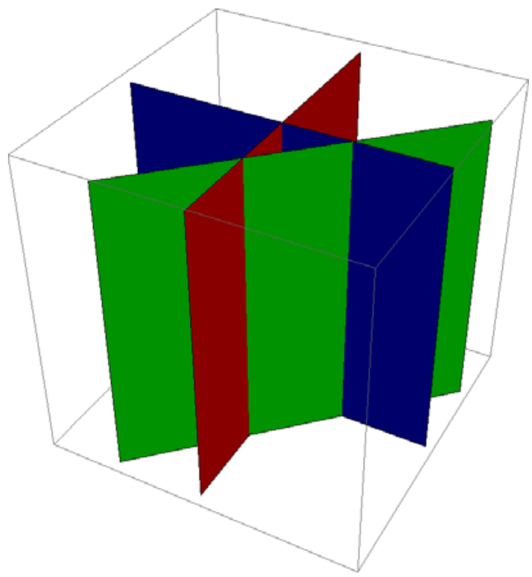


Fig. 2.4. Even without two planes being parallel, it is possible that there is no intersection between all three of them.

BASIC FACT: A system of linear equations has either

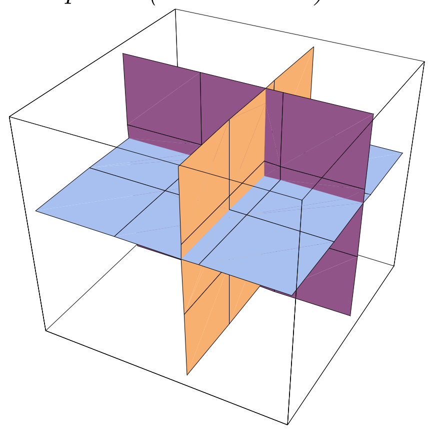
i) exactly one solution *(consistent)* or

ii) infinitely many solutions *(consistent)* or

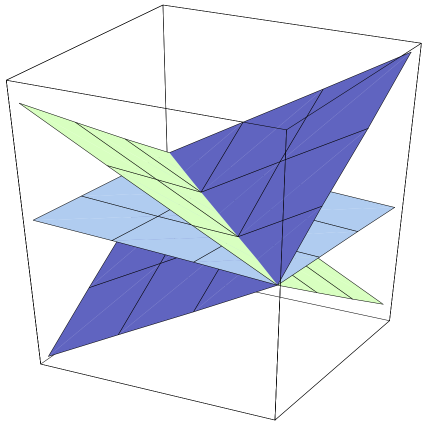
iii) no solution *(inconsistent)*.

Example 2.3 *Three equations in three variables. Each equation determines a plane in 3-space*

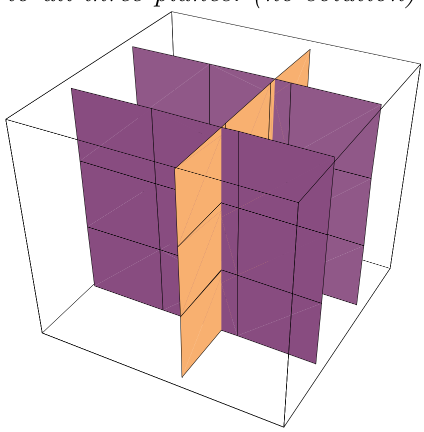
i) The planes intersect in one point. (one solution)



ii) The planes intersect in one line. (infinitely many solutions)



iii) There is not point in common to all three planes. (no solution)



:

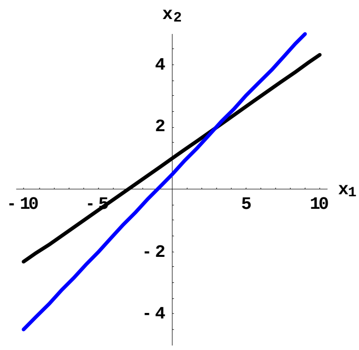
- Two linear systems with the same solution set.

2.1 Strategy for solving a system

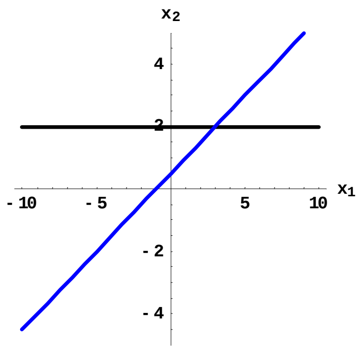
- *Replace one system with an equivalent system that is easier to solve.*

Example 2.4

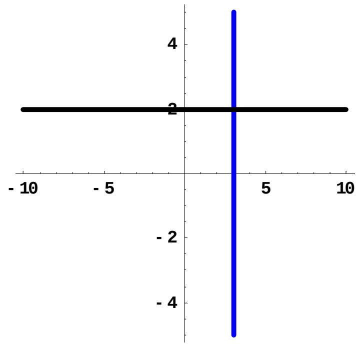
$$\begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ -x_1 & + & 3x_2 & = & 3 \end{array} \longrightarrow \begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ & & x_2 & = & 2 \end{array} \longrightarrow \begin{array}{rcrcr} x_1 & & & = & 3 \\ & & x_2 & = & 2 \end{array}$$



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



$$\begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ & & x_2 & = & 2 \end{array}$$



$$\begin{array}{rcrcr} x_1 & & & = & 3 \\ & & x_2 & = & 2 \end{array}$$

Matrix Notation

$$\begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ -x_1 & + & 3x_2 & = & 3 \end{array} \qquad \underset{\text{(coefficient matrix)}}{\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}}$$

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

$$\downarrow \qquad \text{(augmented matrix)}$$

$$\begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ & & x_2 & = & 2 \end{array} \qquad \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

$$\downarrow$$

$$\begin{array}{rcr} x_1 & = & 3 \\ x_2 & = & 2 \end{array} \qquad \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \end{bmatrix}$$

Elementary Row Operations:

1. *(Replacement)* Add one row to a multiple of another row.
2. *(Interchange)* Interchange two rows.
3. *(Scaling)* Multiply all entries in a row by a nonzero constant.

Each of those three operations has a restriction. Multiplying a row by 0 is not allowed because obviously that can change the solution set of the system. Similarly, adding a multiple of a row to itself is not allowed because adding -1 times the row to itself has the effect of multiplying the row by 0. Finally, swapping (i.e. interchange) a row with itself is disallowed to make some results in the next lectures easier to state and remember (and besides, self-swapping doesn't accomplish anything). The three elementary row operations are sometimes called the *Gaussian operations.* **??**

Row equivalent matrices: Two matrices where one matrix can be transformed into the other matrix by a sequence of elementary row operations.

Fact about Row Equivalence: If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ -4x_1 & + & 5x_2 & + & 9x_3 & = & -9 \end{array} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ & - & 3x_2 & + & 13x_3 & = & -9 \end{array} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & - & 4x_3 & = & 4 \\ & - & 3x_2 & + & 13x_3 & = & -9 \end{array} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & - & 4x_3 & = & 4 \\ & & & & x_3 & = & 3 \end{array} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & & & = & -3 \\ & & x_2 & & & = & 16 \\ & & & & x_3 & = & 3 \end{array} \qquad \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

$$\begin{array}{rcrcrcr} x_1 & & & & & = & 29 \\ & & x_2 & & & = & 16 \\ & & & & x_3 & = & 3 \end{array} \qquad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Solution: $(29, 16, 3)$
Check: Is $(29, 16, 3)$ a solution of the ***original*** system?

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ -4x_1 & + & 5x_2 & + & 9x_3 & = & -9 \end{array}$$

$$\begin{array}{lll} (29) - 2(16) + 3 & = 29 - 32 + 3 & = 0 \\ 2(16) - 8(3) & = 32 - 24 & = 8 \\ -4(29) + 5(16) + 9(3) & = -116 + 80 + 27 & = -9 \end{array}$$

Matrix "jargon". A rectangular array of numbers is called a matrix. If the matrix has m rows and n columns, it is called a $m \times n$ matrix. A matrix with one column only is called a **column vector**, a matrix with one row a **row vector**. The entries of a matrix are denoted by a_{ij} , where i is the row and j is the column.

In the case of the linear equation above, the coefficient matrix A is a square matrix and the augmented matrix B above is a 3×4 matrix.

Row picture and column picture of the same matrix:

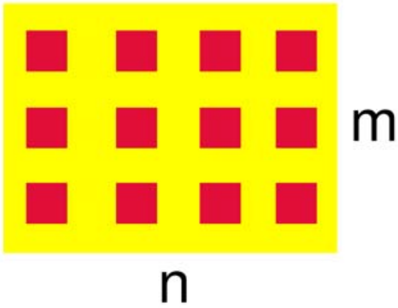


Fig. 2.5. 3×4 matrix.

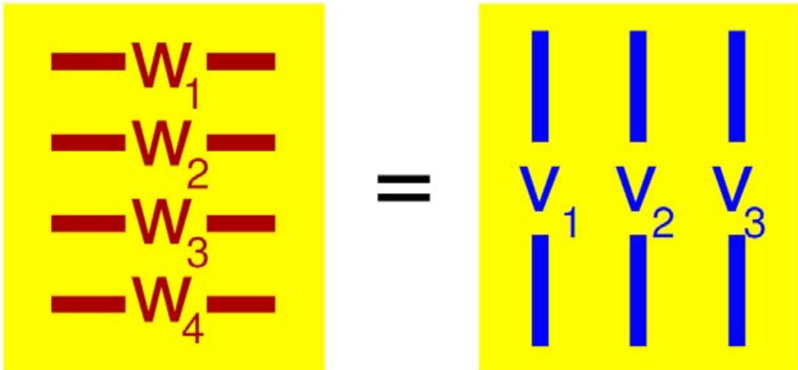


Fig. 2.6. This 4×3 matrix can be viewed as an ordered set of 4 rows or as an ordered set of 3 columns.

Two Fundamental Questions (Existence and Uniqueness)

1. Is the system consistent; (i.e. does a solution **exist**?)
2. If a solution exists, is it **unique**? (i.e. is there one & only one solution?)

Example 2.5 *Is this system consistent?*

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ -4x_1 & + & 5x_2 & + & 9x_3 & = & -9 \end{array}$$

In the last example, this system was reduced to the triangular form:

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & - & 4x_3 & = & 4 \\ & & & & x_3 & = & 3 \end{array} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

This is sufficient to see that the system is consistent and unique. Why?

Example 2.6 *Is this system consistent?*

$$\begin{aligned} 3x_2 - \quad 6x_3 &= 8 \\ x_1 - 2x_2 + \quad 3x_3 &= -1 \\ 5x_1 - 7x_2 + 9x_3 &= 0 \end{aligned} \qquad \begin{bmatrix} 0 & 3 & -6 & 8 \\ 1 & -2 & 3 & -1 \\ 5 & -7 & 9 & 0 \end{bmatrix}$$

Solution: Row operations produce:

$$\begin{bmatrix} 0 & 3 & -6 & 8 \\ 1 & -2 & 3 & -1 \\ 5 & -7 & 9 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 3 & -1 \\ 0 & 3 & -6 & 8 \\ 0 & 3 & -6 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 3 & -1 \\ 0 & 3 & -6 & 8 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

Equation notation of triangular form:

$$\begin{array}{ccccccc} x_1 & - & 2x_2 & + & 3x_3 & = & -1 \\ & & 3x_2 & - & 6x_3 & = & 8 \\ & & & & 0x_3 & = & -3 \quad \leftarrow \textit{Never true} \end{array}$$

The original system is inconsistent!

Example 2.7 *For what values of h will the following system be consistent?*

$$\begin{array}{ccccc} 3x_1 & - & 9x_2 & = & 4 \\ -2x_1 & + & 6x_2 & = & h \end{array}$$

Solution: Reduce to triangular form.

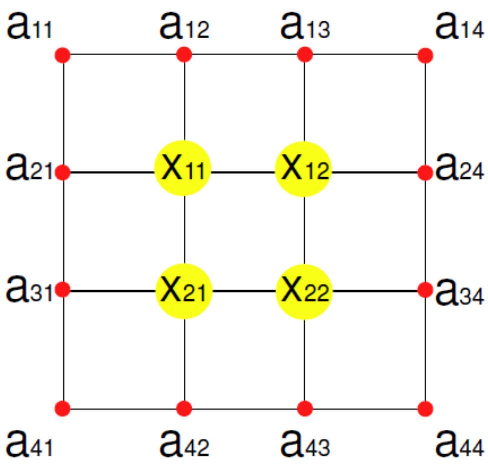
$$\begin{bmatrix} 3 & -9 & 4 \\ -2 & 6 & h \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & \frac{4}{3} \\ -2 & 6 & h \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & \frac{4}{3} \\ 0 & 0 & h + \frac{8}{3} \end{bmatrix}$$

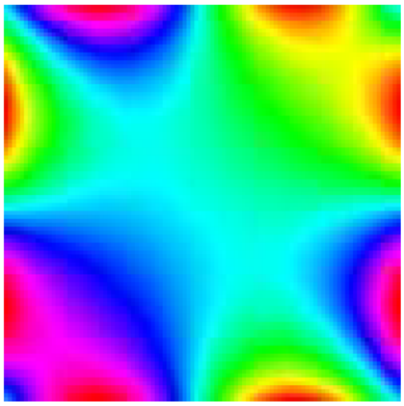
The second equation is $0x_1 + 0x_2 = h + \frac{8}{3}$. System is consistent only if $h + \frac{8}{3} = 0$ or $h = \frac{-8}{3}$.

Example 2.8 *As an example of a system with many variables, consider a drum modeled by a fine net. The heights at each interior node needs the average the heights of the 4 neighboring nodes. The height at the boundary is fixed. With n^2 nodes in the interior, we have to solve a system of n^2 equations. For example, for $n = 2$ (see left), the $n^2 = 4$ equations are*

$$\begin{aligned} x_{11} &= a_{21} + a_{12} + x_{21} + x_{12}, \\ x_{12} &= x_{11} + x_{13} + x_{22} + x_{22}, \\ x_{21} &= x_{31} + x_{11} + x_{22} + a_{43}, \\ x_{22} &= x_{12} + x_{21} + a_{43} + a_{34}. \end{aligned}$$

To the right, we see the solution to a problem with $n = 300$, where the computer had to solve a system with 90000 variables.





Exercise 2.1 *Emile and Gertrude are brother and sister. Emile has twice as many sisters as brothers, and Gertrude has just as many brothers as sisters. How many children are there in this family?*

Exercise 2.2 *On your next trip to Switzerland, you should take the scenic boat ride from Rheinfall to Rheinau and back. The trip downstream from Rheinfall to Rheinau takes* 20 *minutes, and the return trip takes* 40 *minutes; the distance between Rheinfall and Rheinau along the river is* 8 *kilometers. How fast does the boat travel (relative to the water), and how fast does the river Rhein flow in this area? You may assume both speeds to be constant throughout the journey.*

Exercise 2.3 *In a grid of wires, the temperature at exterior mesh points is maintained at constant values (in °C) as shown in the accompanying figure 2.7. When the grid is in thermal equilibrium, the temperature T at each interior mesh point is the average of the temperatures at the four adjacent points. For example,*

$$T_2 = \frac{T_1 + T_3 + 200 + 0}{4}$$

Find the temperatures T_1, T_3, and T_3 when the grid is in thermal equilibrium.

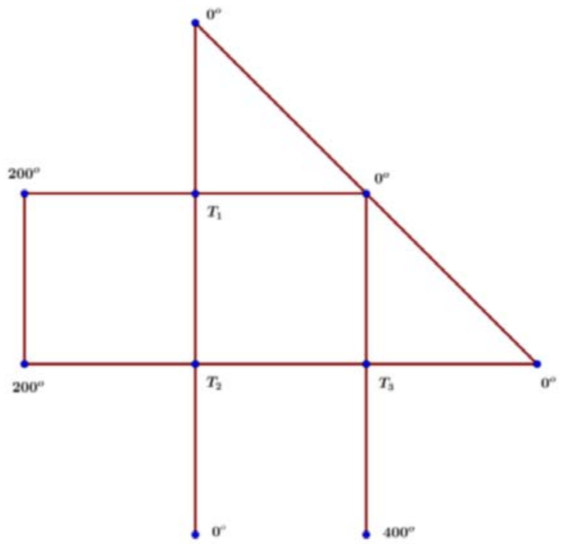


Fig. 2.7. Grid of wires.

3

Row Reduction and Echelon Forms

Echelon form (or **row echelon form**):

1. All nonzero rows are above any rows of all zeros.

2. Each *leading entry* (i.e. left most nonzero entry) of a row is in a column to the right of the leading entry of the row above it.

3. All entries in a column below a leading entry are zero.

Example 3.1 *Echelon forms*

$$\text{(a)} \begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{(b)} \begin{bmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \\ 0 & 0 & 0 \end{bmatrix}$$

$$\text{(c)} \begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \end{bmatrix}$$

Reduced echelon form: Add the following conditions to conditions 1, 2, and 3 above:

4. The leading entry in each nonzero row is 1.

5. Each leading 1 is the only nonzero entry in its column.

Pro memoriam:

Leaders like to be number one, are lonely
and want other leaders above to their left.

Example 3.2 *(continued):*

Reduced echelon form (**rref**) :

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

Theorem 3.3 ***(Uniqueness of The Reduced Echelon Form):*** *Each matrix is row-equivalent to one and only one reduced echelon matrix.*

Important Terms:

- **pivot position:** a position of a leading entry in an echelon form of the matrix.
- **pivot:** a nonzero number that either is used in a pivot position to create 0's or is changed into a leading 1, which in turn is used to create 0's.
- **pivot column:** a column that contains a pivot position.

Example 3.4 *Row reduce to echelon form and locate the pivot columns.*

$$\begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

Solution

pivot
$\swarrow$

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

$\uparrow$
pivot column

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

Possible Pivots:

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{array}{rc} \text{Original Matrix:} & \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \\ \text{pivot columns:} & \begin{matrix} \uparrow & \uparrow & & \uparrow & \\ 1 & 2 & & 4 & \end{matrix} \end{array}$$

Note: There is no more than one pivot in any row. There is no more than one pivot in any column.

Example 3.5 *Row reduce to echelon form and then to reduced echelon form:*

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

Solution:

Step 1: Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \sim \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

Step 2: Use row replacement operations to create zeros in all positions below the pivot.

$$\sim \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

Step 3: Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1–3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \sim \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

$$\sim \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \qquad \text{(echelon form)}$$

Final step to create the reduced echelon form: Beginning with the rightmost leading entry, and working upwards to the left, create zeros above each leading entry and scale rows to transform each leading entry into 1.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \sim \begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \sim$$
$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

3.1 Solutions of linear systems

- **basic variable:** any variable that corresponds to a pivot column in the augmented matrix of a system.
- **free variable:** all nonbasic variables.

SOLUTIONS OF LINEAR EQUATIONS. A system $Ax = b$ with m equations and n unknowns is defined by the $m \times n$ coefficient matrix A and the RHS vector b. The row reduced matrix $\mathrm{rref}(B)$ of the augmented matrix B determines the number of solutions of the system $Ax = b$. There are three possibilities (see Figure 3.1):

- **Consistent:** Exactly one solution. There is a leading 1 in each row but none in the last column of B.
- **Inconsistent:** No solutions. There is a leading 1 in the last column of B.
- **Infinitely many solutions.** There are rows of B without leading 1.

If $m < n$ (less equations then unknowns), then there are either zero or infinitely many solutions.

The $\mathrm{rank}(A)$ of a matrix A is the number of leading ones in $\mathrm{rref}(A)$.

How do we determine in which case we are? It is the rank of A and the rank of the augmented matrix $B = [A|b]$ as well as the number m of columns which determine everything:

If $\text{rank}(A) = \text{rank}(B) = m$: there is exactly 1 solution.

If $\text{rank}(A) < \text{rank}(B)$: there are no solutions.

If $\text{rank}(A) = \text{rank}(B) < m$: there are many solutions.

Example 3.6

$$\begin{bmatrix} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -8 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \quad \begin{array}{ccccccc} x_1 & +6x_2 & & +3x_4 & & = 0 & \text{pivot columns: } 1, 3, 5 \\ & & x_3 & -8x_4 & & = 5 & \text{basic variables: } x_1, x_3, x_5 \\ & & & & x_5 & = 7 & \text{free variables: } x_2, x_4 \end{array}$$

Final Step in Solving a Consistent Linear System: After the augmented matrix is in **reduced** echelon form and the system is written down as a set of equations:

Solve each equation for the basic variable in terms of the free variables (if any) in the equation.

Example 3.7

$$\begin{array}{cccccc} x_1 & +6x_2 & & +3x_4 & & = 0 \\ & & x_3 & -8x_4 & & = 5 \\ & & & & x_5 & = 7 \end{array} \qquad \left\{ \begin{array}{l} x_1 = -6x_2 - 3x_4 \\ x_2 \textit{ is free} \\ x_3 = 5 + 8x_4 \\ x_4 \textit{ is free} \\ x_5 = 7 \\ \textit{(general solution)} \end{array} \right.$$

Example 3.8 (Ex. 3.4 cont.) *The system of linear equations associated with the augmented matrix*

$$B = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \tag{3.1}$$

(see Ex. 3.4) has the following form:

$$\begin{array}{ccccccc} & -3x_2 & -6x_3 & +4x_4 & = & 9 \\ -x_1 & -2x_2 & -1x_3 & +3x_4 & = & 1 \\ -2x_1 & -3x_2 & & +3x_4 & = & -1 \\ x_1 & +4x_2 & +5x_3 & -9x_4 & = & -7 \end{array}. \tag{3.2}$$

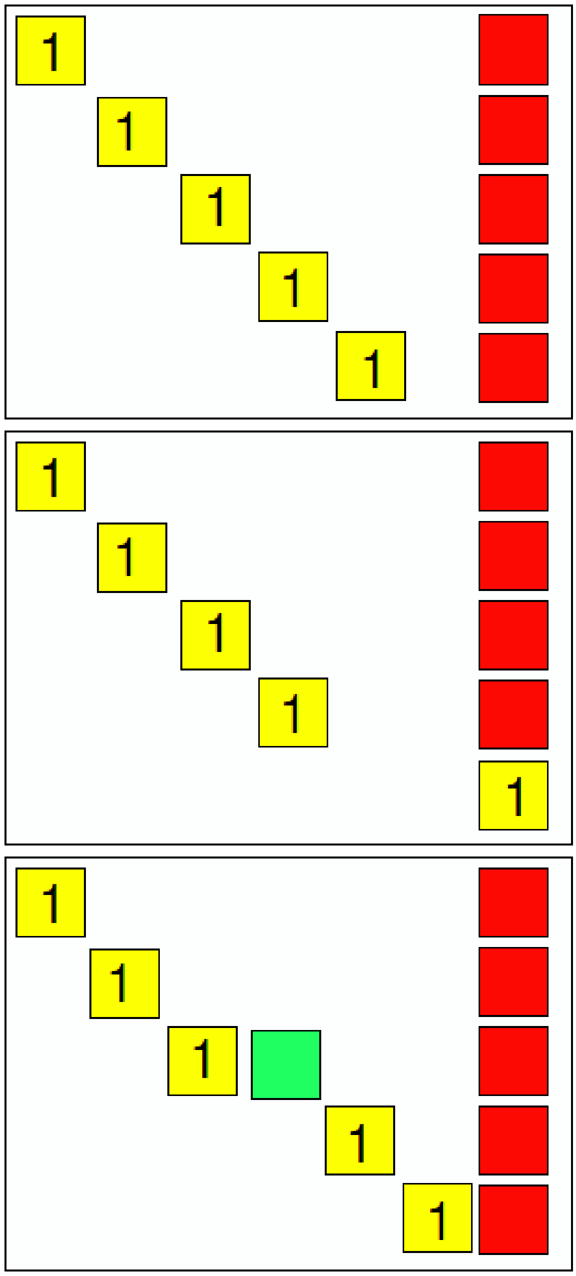


Fig. 3.1. There are three possibilities in the system $Ax = b$.

As the reduced row echelon form of B is equal to

$$\left[\begin{array}{cccc|c} 1 & 0 & -3 & 0 & 5 \\ 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{3.3}$$

(check it!) we conclude, that the general solution of our system is:

$$\begin{array}{rcl} x_1 & = & 5+3t \\ x_2 & = & -3-2t \\ x_3 & = & t \\ x_4 & = & 0 \end{array} \qquad \textit{where values of } t \textit{ are arbitrary.} \qquad \square$$

The **general solution** of the system provides a parametric description of the solution set. (The free variables act as parameters.) The above system has **infinitely many solutions.** Why?

Warning: Use only the reduced echelon form to solve a system.

3.2 Linear Systems with free variables (cont.)

Using Gaussian elimination we can try to solve systems of linear equations with any number of equations and unknowns. We will now look at an example of a linear system with four equations in five unknowns:

$$\left\{\begin{array}{rcrcrcrcrcl} x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 3 \\ 2x_1 & + & x_2 & + & x_3 & + & x_4 & + & 2x_5 & = & 4 \\ x_1 & - & x_2 & - & x_3 & + & x_4 & + & x_5 & = & 5 \\ x_1 & & & & & + & x_4 & + & x_5 & = & 4 \end{array}\right.$$

The augmented matrix is

$$(A|\mathbf{b}) = \left[\begin{array}{cccccc} 1 & 1 & 1 & 1 & 1 & 3 \\ 2 & 1 & 1 & 1 & 2 & 4 \\ 1 & -1 & -1 & 1 & 1 & 5 \\ 1 & 0 & 0 & 1 & 1 & 4 \end{array}\right]$$

Check that your augmented matrix is correct before you proceed, or you could be solving the wrong system! A good method is to first write down the coefficients by rows, reading across the equations, and then to check the columns do correspond to the coefficients of that variable.

Now follow the algorithm to put $(A|\mathbf{b})$ into reduced row echelon form:

$$\begin{matrix} \rightarrow \\ R_2 - 2R_1 \\ R_3 - R_1 \\ R_4 - R_1 \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & -1 & -1 & -1 & 0 & -2 \\ 0 & -2 & -2 & 0 & 0 & 2 \\ 0 & -1 & -1 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{matrix} \\ (-1)\,R_2 \\ \rightarrow \\ \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & -2 & -2 & 0 & 0 & 2 \\ 0 & -1 & -1 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{matrix} \\ \rightarrow \\ R_3 + 2R_2 \\ R_4 + R_2 \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 2 & 0 & 6 \\ 0 & 0 & 0 & 1 & 0 & 3 \end{bmatrix}$$

$$\begin{matrix} \\ \rightarrow \\ \left(\frac{1}{2}\right) R_3 \\ \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 3 \end{bmatrix}$$

$$\begin{matrix} \\ \rightarrow \\ \\ R_4 - R_3 \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This matrix is in row echelon form. We continue to reduced row echelon form, starting with the third row:

$$\begin{matrix} R_1 - R_3 \\ R_2 - R_3 \\ \rightarrow \\ \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{matrix} R_1 - R_2 \\ \\ \rightarrow \\ \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

There are only three leading ones in the reduced row echelon form of this matrix. These appear in columns $1, 2$ and 4. Since the last row gives no information, but merely states that $0 = 0$, the matrix is equivalent to the system of equations:

$$\begin{cases} x_1 & & & & + & x_5 & = & 1 \\ & x_2 & + & x_3 & & & = & -1 \\ & & & & x_4 & & = & 3 \end{cases}$$

The form of these equations tells us that we can assign any values to x_3 and x_5, and then the values of x_1, x_2 and x_4 will be determined.

Definition 3.9 (Leading variables) *The variables corresponding to the columns with leading ones in the reduced row echelon form of an augmented matrix are called* leading variables*. The other variables are called* non-leading variables*.*

In this example, the variables x_1, x_2 and x_4 are leading variables, x_3 and x_5 are non-leading variables. We assign x_3, x_5 the arbitrary values s, t, where s, t represent any real numbers, and then solve for the leading variables in terms of these. We get

$$x_4 = 3, \qquad x_2 = -1 - s, \qquad , x_1 = 1 - t.$$

Then we express this solution in vector form:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1-t \\ -1-s \\ s \\ 3 \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 3 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Observe that this system is consistent and there are infinitely many solutions, because any values of $s \in \mathbb{R}$ and $t \in \mathbb{R}$ will give a solution.

The solution given above is called a *general solution of the system*, because it gives a solution for any values of s and t, and any solution of the equation is of this form for some $s, t \in \mathbb{R}$. For any particular assignment of values to s and t, such as $s = 0$, $t = 1$, we obtain a *particular solution of the system.*

With practice, you will be able to read the general solution directly from the reduced row echelon form of the augmented matrix. We have

$$(A|\mathbf{b}) \to \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Locate the leading ones, and note which are the leading variables. Then locate the non-leading variables and assign each an arbitrary parameter.

So, as above, we note that the leading ones are in the first, second and fourth column, and so correspond to x_1, x_2 and x_4. Then we assign arbitrary parameters to the non-leading variables; that is, values such as $x_3 = s$ and $x_5 = t$, where s and t represent any real numbers. Then write down the vector $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)^T$ (as a column) and fill in the values starting with x_5 and working up. We have $x_5 = t$. Then the third row tells us that $x_4 = 3$. We have $x_3 = s$. Now look at the second row, which says $x_2 + x_3 = -1$, or $x_2 = -1 - s$. Then the top row tells us that $x_1 = 1 - t$. In this way, we obtain the solution in vector form.

3.3 Existence and Uniqueness Questions

Example 3.10

$$\left\{\begin{array}{rrrrrcr} & 3x_2 & -6x_3 & +6x_4 & +4x_5 & = & -5 \\ 3x_1 & -7x_2 & +8x_3 & -5x_4 & +8x_5 & = & 9 \\ 3x_1 & -9x_2 & +12x_3 & -9x_4 & +6x_5 & = & 15 \end{array}\right.$$

In an earlier example, we obtained the echelon form:

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad (x_5 = 4)$$

No equation of the form $0 = c$, where $c \neq 0$, so the system is consistent.
Free variables: x_3 and x_4

$$\boxed{\textbf{Consistent system with free variables}} \implies \textbf{infinitely many solutions.}$$

$$\begin{array}{rrcr} 3x_1 & +4x_2 & = & -3 \\ 2x_1 & +5x_2 & = & 5 \\ -2x_1 & -3x_2 & = & 1 \end{array} \rightarrow \begin{bmatrix} 3 & 4 & -3 \\ 2 & 5 & 5 \\ -2 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 3 & 4 & -3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow$$

$$3x_1 + 4x_2 = -3$$
$$x_2 = 3$$

$$\boxed{\textbf{Consistent system, no free variables}} \implies \textbf{unique solution.}$$

Theorem 3.11 ***(Existence and Uniqueness Theorem)***

1. *A linear system is consistent if and only if the rightmost column of the augmented matrix is not a pivot column, i.e., if and only if an echelon form of the augmented matrix has no row of the form*

$$\begin{bmatrix} 0 & \dots & 0 & b \end{bmatrix} \quad (\textit{where } b \textit{ is nonzero}).$$

2. *If a linear system is consistent, then the solution contains either*

   i) *a unique solution (when there are no free variables) or*

   ii) *infinitely many solutions (when there is at least one free variable).*

Using Row Reduction to Solve Linear Systems

1. Write the augmented matrix of the system.
2. Use the row reduction algorithm to obtain an equivalent augmented matrix in echelon form. Decide whether the system is consistent. If not, stop; otherwise go to the next step.
3. Continue row reduction to obtain the reduced echelon form.
4. Write the system of equations corresponding to the matrix obtained in step 3.
5. State the solution by expressing each basic variable in terms of the free variables and declare the free variables.

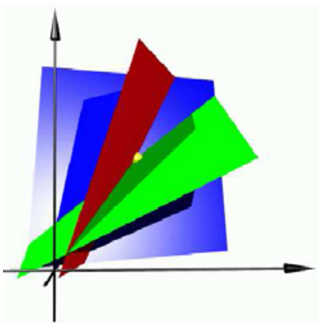
1. **a)** What is the largest possible number of pivots a 4×6 matrix can have? Why?

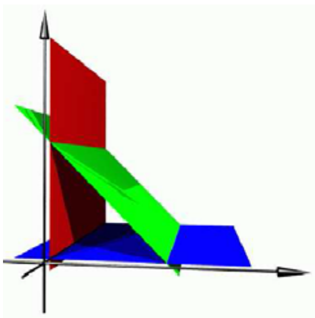
   b) What is the largest possible number of pivots a 6×4 matrix can have? Why?

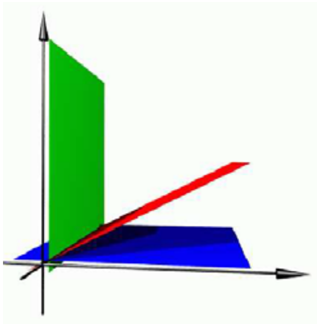
   c) How many solutions does a consistent linear system of 3 equations and 4 unknowns have? Why?

   d) Suppose the coefficient matrix corresponding to a linear system is 4×6 and has 3 pivot columns. How many pivot columns does the augmented matrix have if the linear system is inconsistent?

Exercise 3.1 *The reduced echelon form of the augmented matrix B determines on how many solutions the linear system $Ax = b$ has. Check, if the answers are correct:*

The "good"	1 solution
	$\left[\begin{array}{rrr\|r} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 2 & 1 & -1 & -2 \end{array}\right]$ $\left[\begin{array}{rrr\|r} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 2 & 1 & -1 & -2 \end{array}\right]$ $\left[\begin{array}{rrr\|r} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 3 & -3 & -12 \end{array}\right]$ $\left[\begin{array}{rrr\|r} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & -1 & -4 \end{array}\right]$ $\left[\begin{array}{rrr\|r} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & -3 & -6 \end{array}\right]$ $\left[\begin{array}{rrr\|r} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 2 \end{array}\right]$ $\left[\begin{array}{rrr\|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{array}\right]$ $\text{rank}(A) = 3 \quad \text{rank}(B) = 3$

The "bad"	0 solution
	$\left[\begin{array}{ccc\|c} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 1 & 0 & 3 & -2 \end{array}\right]$ $\left[\begin{array}{ccc\|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 1 & 0 & 3 & -2 \end{array}\right]$ $\left[\begin{array}{ccc\|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & -7 \end{array}\right]$ $\left[\begin{array}{ccc\|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & -9 \end{array}\right]$ $\left[\begin{array}{ccc\|c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & -9 \end{array}\right]$ $\left[\begin{array}{ccc\|c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 \end{array}\right]$ $\text{rank}(A) = 2 \quad \text{rank}(B) = 3$

The "ugly"	∞ solutions
	$\left[\begin{array}{ccc\|c} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 1 & 0 & 3 & 7 \end{array}\right]$
	$\left[\begin{array}{ccc\|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 1 & 0 & 3 & 7 \end{array}\right]$
	$\left[\begin{array}{ccc\|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{array}\right]$
	$\left[\begin{array}{ccc\|c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$
	$\left[\begin{array}{ccc\|c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$
	$\text{rank}(A) = 2 \quad \text{rank}(B) = 2$

Exercise 3.2 (Interpolation) *Find the equation of the parabola which passes through the points* $P = (0; -1)$, $Q = (1; 4)$ *and* $R = (2; 13)$.

Exercise 3.3 15 *kids have bicycles or tricycles. Together they count* 37 *wheels. How many have bicycles?"*

Exercise 3.4 *This system is not linear, in some sense,*

$$
\begin{aligned}
2\sin\alpha - \cos\beta + 3\tan\gamma &= 3 \\
4\sin\alpha + 2\cos\beta - 2\tan\gamma &= 10 \\
6\sin\alpha - 3\cos\beta + \tan\gamma &= 9
\end{aligned}
$$

and yet we can nonetheless apply Gauss' method. Do so. Does the system have a solution?

4

Vector equations

Important properties of linear systems can be described with the concept and notation of vectors. This chapter connects equations involving vectors to ordinary systems of equations.

Key concepts to master: linear combinations of vectors and a spanning set.

Vector: A matrix with only one column.

Vectors in $\mathbb{R}^n$ (vectors with n entries):

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

4.1 Geometric Description of $\mathbb{R}^2$

Vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is the point (x_1, x_2) in the plane.
$\mathbb{R}^2$ is the set of all points in the plane.

Two vectors in $\mathbb{R}^2$ are **equal** if and only if their corresponding entries are equal. Thus $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ are not equal, because vectors in $\mathbb{R}^2$ are ordered pairs of real numbers.

Given two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$, their sum is the vector $\mathbf{u} + \mathbf{v}$ obtained by adding corresponding entries of $\mathbf{u}$ and $\mathbf{v}$. For example,

$$\begin{bmatrix} 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 5 \\ 2 \end{bmatrix} =: \begin{bmatrix} 7 \\ 7 \end{bmatrix}.$$

Given a vector **u** and a real number c, the scalar multiple of **u** by c is the vector $c\mathbf{u}$ obtained by multiplying each entry in **u** by c. For instance,

$$\text{if} \quad \mathbf{u} = \begin{bmatrix} 2 \\ 5 \end{bmatrix} \quad \text{and} \quad c = 3, \qquad \text{then } c\mathbf{u} = 3 \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 6 \\ 15 \end{bmatrix}.$$

The number c in $c\mathbf{u}$ is called a *scalar*; it is written in lightface type to distinguish it from the boldface vector **u**.

The operations of scalar multiplication and vector addition can be combined, as in the following example.

Example 4.1 *Given vector* $\mathbf{u} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ *and* $\mathbf{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, *find* $(-3)\,\mathbf{u} + 2\mathbf{v}$.

Solution:

$$(-3)\,\mathbf{u} = \begin{bmatrix} -6 \\ -15 \end{bmatrix}, \qquad 2\mathbf{v} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

and

$$(-3)\,\mathbf{u} + 2\mathbf{v} = \begin{bmatrix} -8 \\ -11 \end{bmatrix} \qquad \square$$

Parallelogram rule for addition of two vectors:
If **u** and **v** in $\mathbb{R}^2$ are represented as points in the plane, then $\mathbf{u}+\mathbf{v}$ corresponds to the fourth vertex of the parallelogram whose other vertices are **0**, **u** and **v**. (Note that $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.)

Example 4.2 *Let* $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ *and* $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. *Graphs of* $\mathbf{u}, \mathbf{v}$ *and* $\mathbf{u} + \mathbf{v}$ *are given below (see Fig.4.1):*

Example 4.3 *Let* $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. *Express* **u**, $2\mathbf{u}$, *and* $\frac{-3}{2}\mathbf{u}$ *on a graph (see Fig. 4.2).*

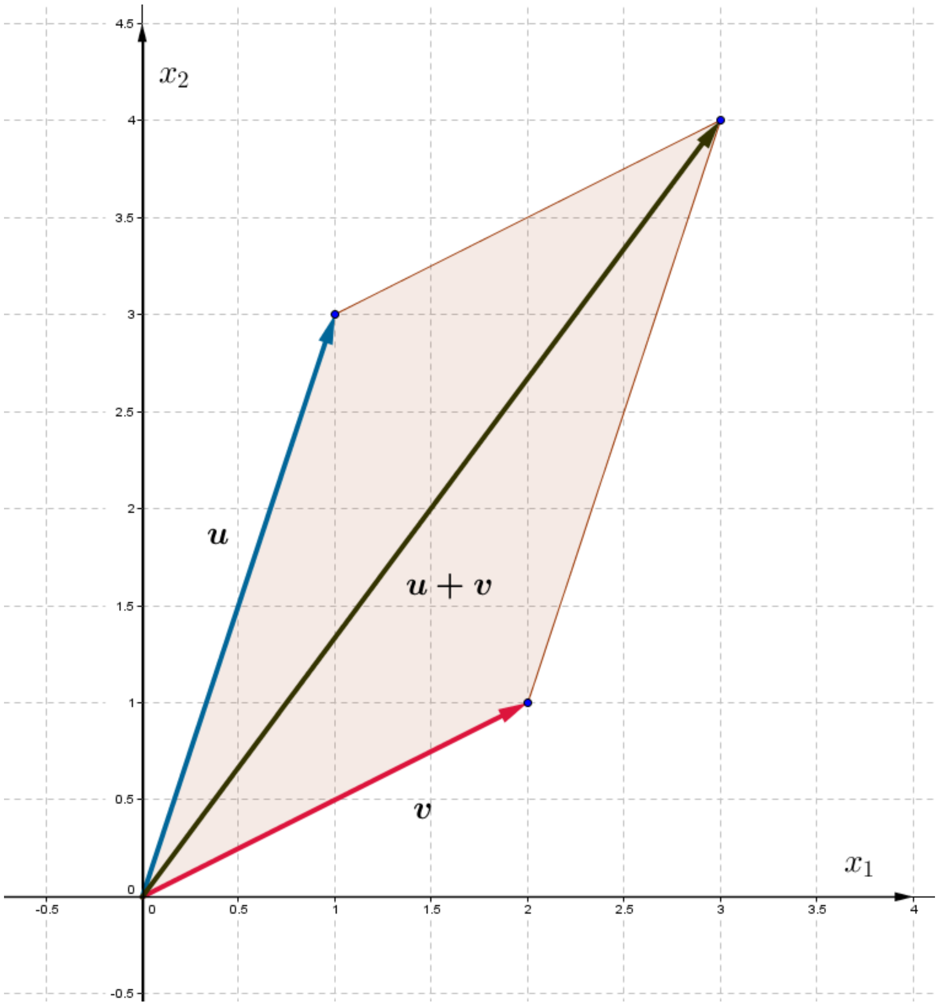


Fig. 4.1. Illustration of the Parallelogram Rule

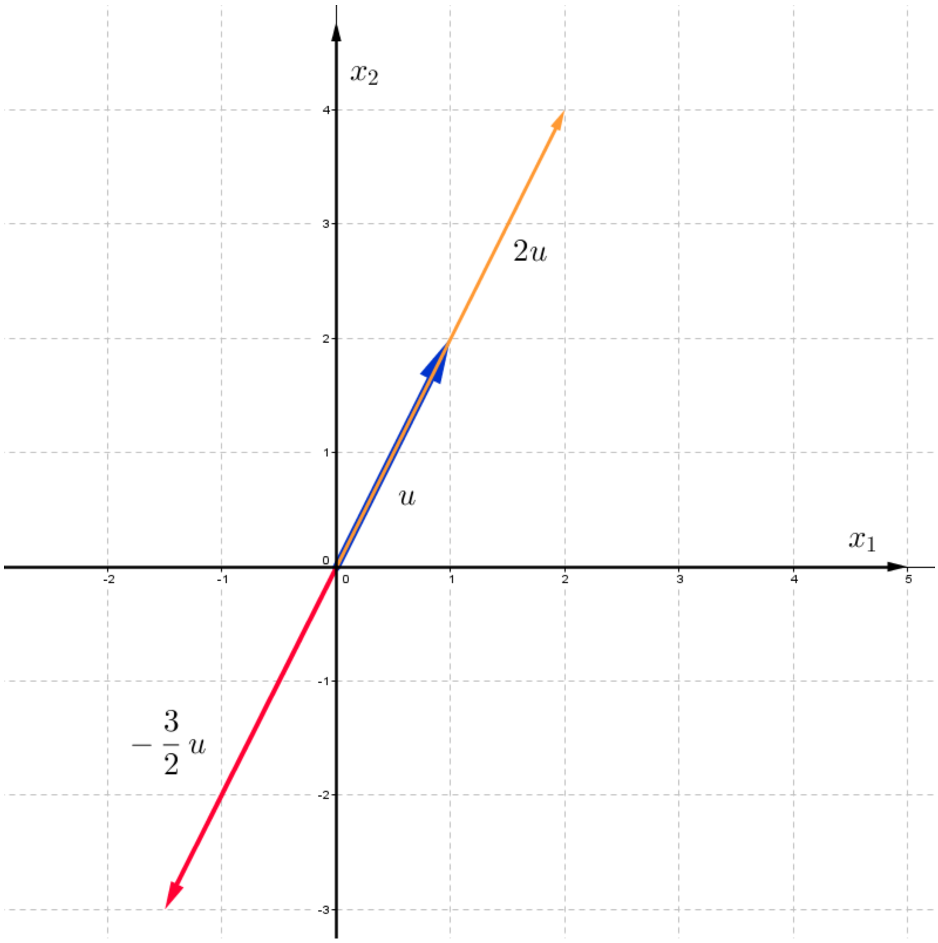


Fig. 4.2. Colinear vectors.

Linear Combinations

Algebraic Properties of $\mathbb{R}^n$:

Operations on vectors have the following properties, which can be verified directly from the corresponding properties for real numbers.

1. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$
2. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$
3. $\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}$
4. $\mathbf{u}+(-\mathbf{u})=-\mathbf{u}+\mathbf{u}=\mathbf{0}$, where $-\mathbf{u}$ denotes $(-1)\mathbf{u}$
5. $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$
6. $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$
7. $c(d\mathbf{u})=(cd)\mathbf{u}$
8. $1\mathbf{u}=\mathbf{u}$

Definition 4.4 *Given vectors $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1,c_2,\ldots,c_p$, the vector $\mathbf{y}$ defined by*

$$\mathbf{y}=c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_p\mathbf{v}_p$$

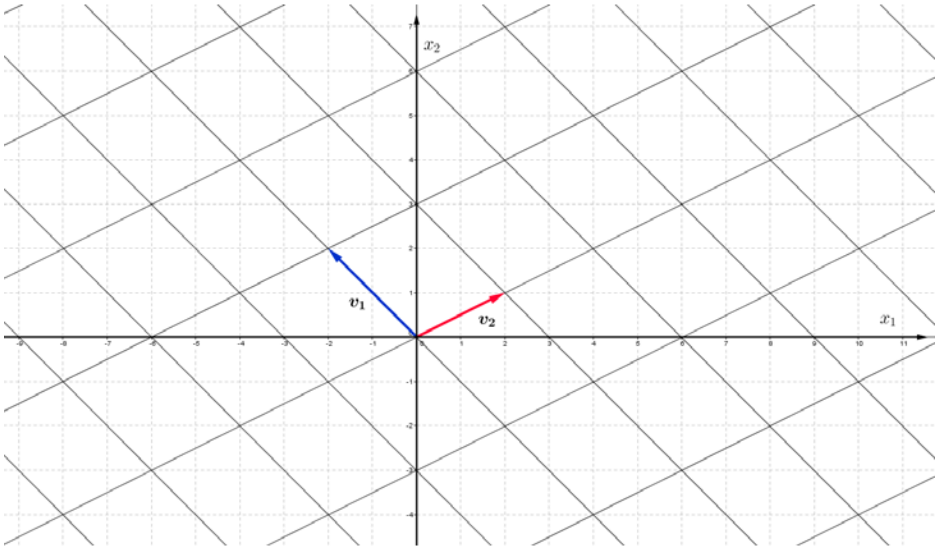
*is called a **linear combination** of $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_p$ using weights $c_1,c_2,\ldots,c_p$.*

Examples of linear combinations of $\mathbf{v}_1$and $\mathbf{v}_2$:

$$3\mathbf{v}_1+2\mathbf{v}_2, \qquad \tfrac{1}{3}\mathbf{v}_1, \qquad \mathbf{v}_1-2\mathbf{v}_2, \qquad \mathbf{0}$$

Example 4.5 *Let $\mathbf{v}_1=\begin{bmatrix}2\\1\end{bmatrix}$ and $\mathbf{v}_2=\begin{bmatrix}-2\\2\end{bmatrix}$. Express each of the following as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$:*

$$\mathbf{a}=\begin{bmatrix}0\\3\end{bmatrix},\ \mathbf{b}=\begin{bmatrix}-4\\1\end{bmatrix},\ \mathbf{c}=\begin{bmatrix}6\\6\end{bmatrix},\ \mathbf{d}=\begin{bmatrix}7\\-4\end{bmatrix}$$



Example 4.6 *Let* $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 4 \\ 2 \\ 14 \end{bmatrix}$, $\mathbf{a}_3 = \begin{bmatrix} 3 \\ 6 \\ 10 \end{bmatrix}$, *and* $\mathbf{b} = \begin{bmatrix} -1 \\ 8 \\ -5 \end{bmatrix}$.

Determine if $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$.

Solution: Vector $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ if can we find weights x_1, x_2, x_3 such that

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

Vector Equation (fill-in):

Corresponding System:

$$\begin{array}{rcrcrcr} x_1 & + & 4x_2 & + & 3x_3 & = & -1 \\ & & 2x_2 & + & 6x_3 & = & 8 \\ 3x_1 & + & 14x_2 & + & 10x_3 & = & -5 \end{array}$$

Corresponding Augmented Matrix:

$$\begin{bmatrix} 1 & 4 & 3 & -1 \\ 0 & 2 & 6 & 8 \\ 3 & 14 & 10 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{bmatrix} \qquad \Longrightarrow \begin{array}{l} x_1 = \_\_\_ \\ x_2 = \_\_\_ \\ x_3 = \_\_\_ \end{array}$$

Review of the last example: $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ and $\mathbf{b}$ are columns of the augmented matrix

$$\begin{bmatrix} 1 & 4 & 3 & -1 \\ 0 & 2 & 6 & 8 \\ 3 & 14 & 10 & -5 \end{bmatrix}$$
$$\begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{b} \end{array}$$

Solution to

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b}$$

is found by solving the linear system whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{b} \end{bmatrix}.$$

A vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

has the same solution set as the linear system whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}.$$

In particular, $\mathbf{b}$ can be generated by a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ if and only if there is a solution to the linear system corresponding to the augmented matrix.

4.2 The Span of a Set of Vectors

Example 4.7 *Let* $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$. *Label the origin* $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ *together with* $\mathbf{v}$, $2\mathbf{v}$ *and* $1.5\mathbf{v}$ *on the graph below (Figure* **??***).*

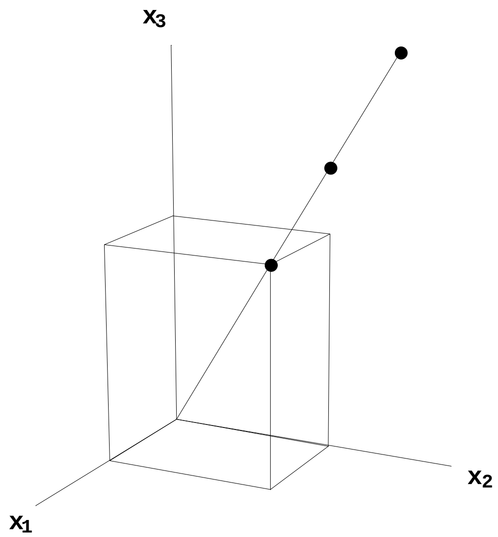


Fig. 4.3. $\mathbf{v}$, $2\mathbf{v}$ and $1.5\mathbf{v}$ all lie on the same line. $\mathbf{Span}\{\mathbf{v}\}$ is the set of all vectors of the form $c\mathbf{v}$. Here, $\mathbf{Span}\{\mathbf{v}\}$ = a line through the origin.

Example 4.8 *Label* $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u}+\mathbf{v}$ *and* $3\mathbf{u}+4\mathbf{v}$ *on the graph below (Figure 4.4).*

.

Definition 4.9 *Suppose* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ *are in* $\mathbb{R}^n$*; then*

Span$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ = *set of all linear combinations of* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$.

Stated another way: $\mathbf{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is the collection of all vectors that can be written as

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p$$

where $x_1, x_2, \ldots, x_p$ are scalars.

Example 4.10 *Let* $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ *and* $\mathbf{v}_2 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$

a) *Find a vector in* ***Span***$\{\mathbf{v}_1, \mathbf{v}_2\}$.

b) *Describe* ***Span***$\{\mathbf{v}_1, \mathbf{v}_2\}$ *geometrically.*

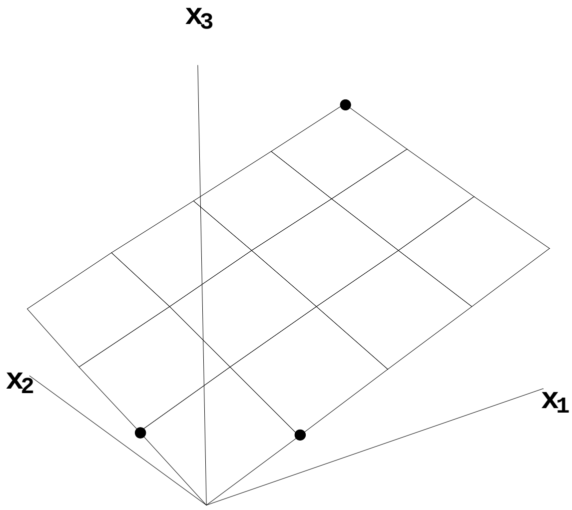


Fig. 4.4. $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u}+\mathbf{v}$ and $3\mathbf{u}$+$4\mathbf{v}$ all lie in the same plane. $\mathbf{Span}\{\mathbf{u},\mathbf{v}\}$ is the set of all vectors of the form $x_1\mathbf{u}+x_2\mathbf{v}$. Here, $\mathbf{Span}\{\mathbf{u},\mathbf{v}\}$ = a plane through the origin.

4.3 Spanning Sets in $\mathbb{R}^3$

Example 4.11 *Let* $\mathbf{v}_1 = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}$ *and* $\mathbf{v}_2 = \begin{bmatrix} 6 \\ 3 \\ 3 \end{bmatrix}$. *Is* ***Span***$\{\mathbf{v}_1,\mathbf{v}_2\}$ *a line or a plane (Figure 4.5)?*

Example 4.12 *Let* $A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 5 \end{bmatrix}$ *and* $\mathbf{b} = \begin{bmatrix} 8 \\ 3 \\ 17 \end{bmatrix}$. *Is* $\mathbf{b}$ *in the plane spanned by the columns of* A*?*

Solution:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 5 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 8 \\ 3 \\ 17 \end{bmatrix}$$

Do x_1 and x_2 exist so that

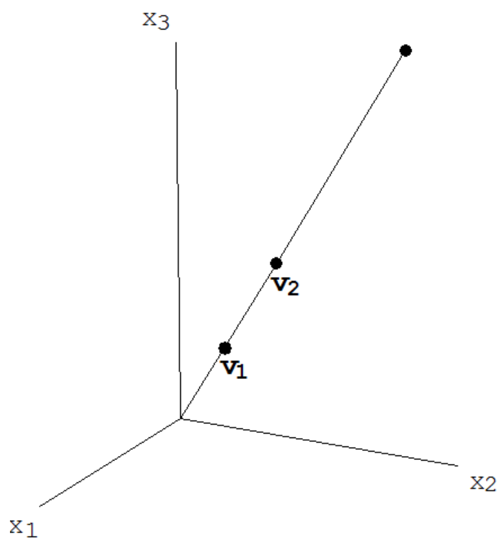


Fig. 4.5. $\mathbf{v}_2$ is a multiple of $\mathbf{v}_1$, $\mathbf{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ =$\mathbf{Span}\{\mathbf{v}_1\}$ =$\mathbf{Span}\{\mathbf{v}_2\}$ (line through the origin)

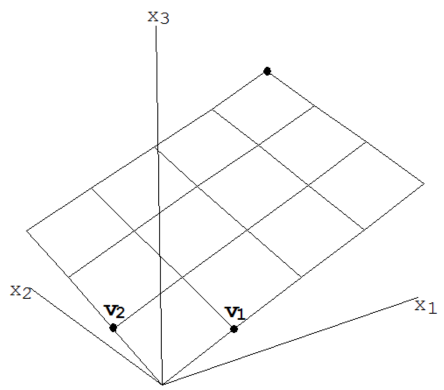


Fig. 4.6. $\mathbf{v}_2$ is **not** a multiple of $\mathbf{v}_1$, $\mathbf{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ =plane through the origin

Corresponding augmented matrix:

$$\begin{bmatrix} 1 & 2 & 8 \\ 3 & 1 & 3 \\ 0 & 5 & 17 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 8 \\ 0 & -5 & -21 \\ 0 & 5 & 17 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 8 \\ 0 & -5 & -21 \\ 0 & 0 & -4 \end{bmatrix}$$

So **b** is not in the plane spanned by the columns of A

5

The Matrix Equation $\mathbf{Ax} = \mathbf{b}$

Linear combinations can be viewed as a matrix-vector multiplication.

Definition 5.1 *If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the **product of** A **and** $\mathbf{x}$, denoted by $A\mathbf{x}$, is the **linear combination of the columns of** $\mathbf{A}$ **using the corresponding entries in** $\mathbf{x}$ **as weights**. I.e.,*

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\,\mathbf{a}_1 + x_2\,\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

$$\begin{bmatrix} 1 & -4 \\ 3 & 2 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 7 \\ -6 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} + (-6) \begin{bmatrix} -4 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 7 \\ 21 \\ 0 \end{bmatrix} + \begin{bmatrix} 24 \\ -12 \\ -30 \end{bmatrix} =$$

$$\begin{bmatrix} 31 \\ 9 \\ -30 \end{bmatrix}$$

Example 5.2 *Write down the system of equations corresponding to the augmented matrix below and then express the system of equations in vector form and finally in the form $A\mathbf{x} = \mathbf{b}$ where $\mathbf{b}$ is a 3×1 vector.*

$$\begin{bmatrix} 2 & 3 & 4 & 9 \\ -3 & 1 & 0 & -2 \end{bmatrix}$$

Solution: Corresponding system of equations (fill-in)

Vector Equation: (fill in)

$$\begin{bmatrix} 2 \\ -3 \end{bmatrix} + \qquad \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \qquad \begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 9 \\ -2 \end{bmatrix}.$$

Matrix equation (fill-in):

Three equivalent ways of viewing a linear system:

1. as a system of linear equations;
2. as a vector equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$**; or**
3. as a matrix equation $A\mathbf{x} = \mathbf{b}$.

Theorem 5.3 *If A is a $m \times n$ matrix, with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and if $\mathbf{b}$ is in $\mathbb{R}^m$, then the matrix equation*

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

- *has the same solution set as the vector equation*

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

- *which, in turn, has the same solution set as the system of linear equations whose augmented matrix is*

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}.$$

Useful Fact:

The equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a

______________ __________ of the columns of A.

Example 5.4 *Let* $A = \begin{bmatrix} 1 & 4 & 5 \\ -3 & -11 & -14 \\ 2 & 8 & 10 \end{bmatrix}$ *and* $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. *Is the equation* $A\mathbf{x} = \mathbf{b}$ *consistent for all* $\mathbf{b}$*?*

Solution: Augmented matrix corresponding to $A\mathbf{x} = \mathbf{b}$:

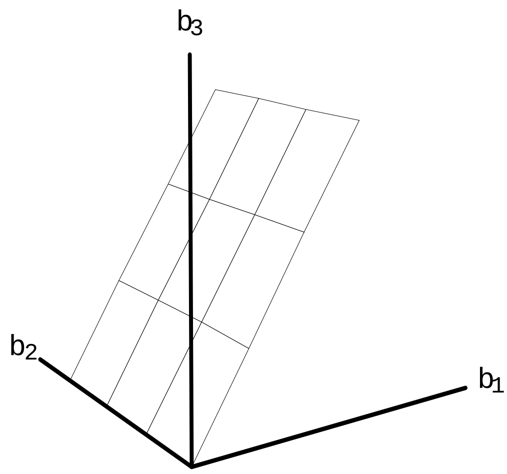
$$\begin{bmatrix} 1 & 4 & 5 & b_1 \\ -3 & -11 & -14 & b_2 \\ 2 & 8 & 10 & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 5 & b_1 \\ 0 & 1 & 1 & 3b_1 + b_2 \\ 0 & 0 & 0 & -2b_1 + b_3 \end{bmatrix}$$

$A\mathbf{x} = \mathbf{b}$ is ________ consistent for all $\mathbf{b}$ since some choices of $\mathbf{b}$ make $-2b_1 + b_3$ nonzero.

$$A = \begin{bmatrix} 1 & 4 & 5 \\ -3 & -11 & -14 \\ 2 & 8 & 10 \end{bmatrix}$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{matrix}$$

The equation $A\mathbf{x} = \mathbf{b}$ is consistent if

$$-2b_1 + b_3 = 0.$$
(equation of a plane in $\mathbb{R}^3$)
$$x_1\mathbf{a}_1 + x_2\mathbf{a}_3 + x_3\mathbf{a}_3 = \mathbf{b} \text{ if and only if } b_3 - 2b_1 = 0.$$



Columns of A span a plane
in $\mathbb{R}^3$ through $\mathbf{0}$

Instead, if *any* $\mathbf{b}$ in $\mathbb{R}^3$ (not just those lying on a particular line or in a plane) can be expressed as a linear combination of the columns of A, then we say that the columns of A span $\mathbb{R}^3$.

Definition 5.5 *We say that **the columns of** $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_p \end{bmatrix}$ **span** $\mathbb{R}^m$ if every vector $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_p$ (i.e. $\text{Span}\,\{\mathbf{a}_1, \ldots, \mathbf{a}_p\} = \mathbb{R}^m$).*

Theorem 5.6 *Let A be an $m \times n$ matrix. Then the following statements are logically equivalent:*

a) *For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.*

b) *Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.*

c) *The columns of A span $\mathbb{R}^m$.*

d) *A has a pivot position in every row.*

Proof. ***(outline)**:* Statements (a), (b) and (c) are logically equivalent.
To complete the proof, we need to show that (a) is true when (d) is true and (a) is false when (d) is false.

Suppose (d) is ______________. Then row-reduce the augmented matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$:

$$[A \quad \mathbf{b}] \sim \cdots \sim [U \quad \mathbf{d}]$$

and each row of U has a pivot position and so there is no pivot in the last column of $[U \quad \mathbf{d}]$.
So (a) is ______________.Now suppose (d) is ______________.
Then the last row of $[U \quad \mathbf{d}]$ contains all zeros.
Suppose $\mathbf{d}$ is a vector with a 1 as the last entry. Then $[U \quad \mathbf{d}]$ represents an inconsistent system.
Row operations are reversible: $[U \quad \mathbf{d}] \sim \cdots \sim [A \quad \mathbf{b}]$

$\Longrightarrow [A \quad \mathbf{b}]$ is inconsistent also. So (a) is _____________. ■

Example 5.7 *Let* $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ *and* $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. *Is the equation* $A\mathbf{x} = \mathbf{b}$ *consistent for all possible* $\mathbf{b}$*?*

Solution: A has only ______ columns and therefore has at most ______ pivots.
Since A does not have a pivot in every __________, $A\mathbf{x} = \mathbf{b}$

is ___________ ___________________________ for all possible $\mathbf{b}$, according to Theorem 5.6.

Example 5.8 *Do the columns of* $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 0 & 3 & 9 \end{bmatrix}$ *span* $\mathbb{R}^3$*?*

Solution:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 0 & 3 & 9 \end{bmatrix} \sim$$

(no pivot in row 2)

By Theorem 5.6, the columns of A _____________.

5.1 Another method for computing $A\mathbf{x}$:

The previous calculations were based on the definition of the product of a matrix A and a vector $\mathbf{x}$. The following simple example will lead to a different (row-oriented) method for calculating the entries in $A\mathbf{x}$ useful, when working problems by hand.

Example 5.9 *Compute* $A\mathbf{x}$*, where*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Solution: From the definition

$$\begin{aligned} & \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ = \; & x_1 \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \qquad (5.1) \\ = \; & \begin{bmatrix} x_1 \\ 4x_1 \\ 7x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 5x_2 \\ 8x_2 \end{bmatrix} + \begin{bmatrix} 3x_3 \\ 6x_3 \\ 9x_3 \end{bmatrix} \\ = \; & \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \\ 7x_1 + 8x_2 + 9x_3 \end{bmatrix} \end{aligned}$$

The first entry in the product $A\mathbf{x}$ is a sum of products (sometimes called a *dot product*), using the first row of A and the entries in $\mathbf{x}$. That is,

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

This matrix shows how to compute the first entry in $A\mathbf{x}$ directly, without writing down all the calculations shown in (5.1). Similarly, the second entry in $A\mathbf{x}$ can be calculated at once by multiplying the entries in the second row of A by the corresponding entries in $\mathbf{x}$ and then summing the resulting products:

$$\begin{bmatrix} & & \\ 4 & 5 & 6 \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \\ 4x_1 + 5x_2 + 6x_3 \end{bmatrix}$$

Likewise, the third entry in $A\mathbf{x}$ can be calculated from the third row of A and the entries in x. □

Theorem 5.10 *If A is an $m \times n$ matrix, $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, and c is a scalar, then:*

a) $A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;

b) $A(c\mathbf{u}) = cA\mathbf{u}$.

6

Solutions Sets of Linear Systems

Solution sets of linear systems are important objects of study in linear algebra. They will appear later in several different contexts. This chapter uses vector notation to give explicit and geometric descriptions of such solution sets.

Homogeneous System:

$$A\mathbf{x} = \mathbf{0}$$

(A is $m \times n$ and $\mathbf{0}$ is the zero vector in $\mathbb{R}^m$)

Example 6.1

$$\begin{array}{rcrcl} x_1 & + & 10x_2 & = & 0 \\ 2x_1 & + & 20x_2 & = & 0 \end{array}$$

Corresponding matrix equation $A\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 10 \\ 2 & 20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Trivial solution:

$$\mathbf{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{or} \qquad \mathbf{x} = \mathbf{0}$$

The homogeneous system $A\mathbf{x} = \mathbf{0}$ *always* has the **trivial solution**, $\mathbf{x} = \mathbf{0}$.

Nonzero vector solutions are called **nontrivial solutions.**
Do **nontrivial** solutions exist?

$$\begin{bmatrix} 1 & 10 & 0 \\ 2 & 20 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 10 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Consistent system with a free variable has infinitely many solutions.

A homogeneous equation $A\mathbf{x} = \mathbf{0}$ has nontrivial solutions if and only if the system of equations **has at least one free variable**.

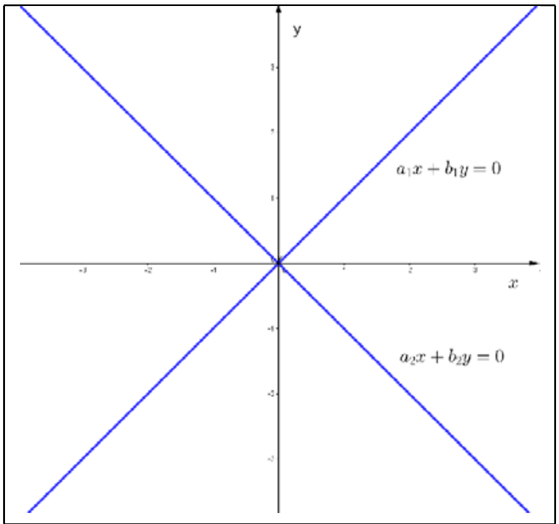
Because a homogeneous linear system always has the trivial solution, **there are only two possibilities for its solutions**:

1. The system has only the trivial solution.
2. The system has infinitely many solutions in addition to the trivial solution.

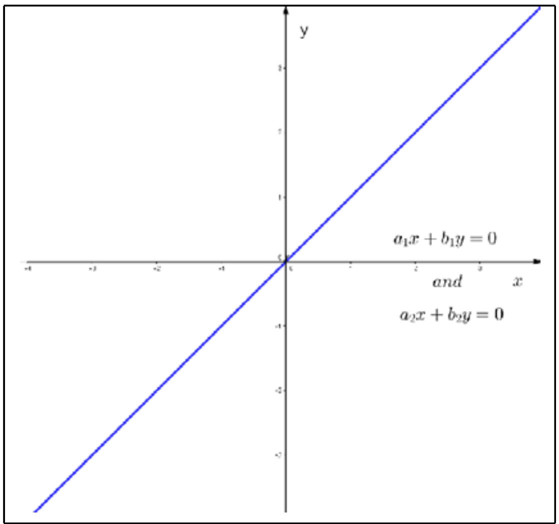
In the special case of a homogeneous linear system of two equations in two unknowns, say

$$\begin{aligned} a_1x + b_1y &= 0 \qquad (a_1,\ b_1 \text{ not both zero}) \\ a_2x + b_2y &= 0 \qquad (a_2,\ b_2 \text{ not both zero}) \end{aligned}$$

the graphs of the equations are lines through the origin, and the trivial solution corresponds to the point of intersection at the origin



a) Only the trivial solution



b) Infinitely many solutions

Example 6.2 *Determine if the following homogeneous system has nontrivial solutions and then describe the solution set.*

$$\begin{array}{rcrcrcl} 2x_1 & + & 4x_2 & - & 6x_3 & = & 0 \\ 4x_1 & + & 8x_2 & - & 10x_3 & = & 0 \end{array}$$

Solution:

There is at least one free variable (why?)

$\Longrightarrow$ nontrivial solutions exist

$$\begin{bmatrix} 2 & 4 & -6 & 0 \\ 4 & 8 & -10 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 0 \\ 4 & 8 & -10 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

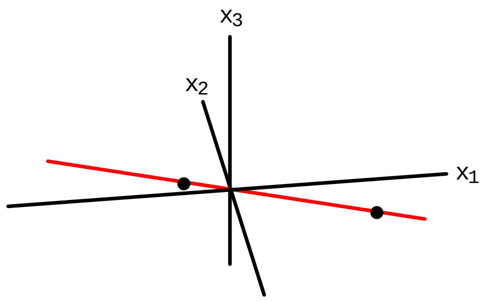
$x_1 =$

x_2 is free

$x_3 =$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_2 \\ x_2 \\ 0 \end{bmatrix} = \_\_\_\_ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = x_2\mathbf{v}$$

Graphical representation:



$$\text{solution set} = \text{span}\{\mathbf{v}\} = \text{line through } \mathbf{0} \text{ in } \mathbb{R}^3$$

Example 6.3 *Describe the solution set of*

$$\begin{array}{rcrcrcr} 2x_1 & + & 4x_2 & - & 6x_3 & = & 0 \\ 4x_1 & + & 8x_2 & - & 10x_3 & = & 4 \end{array}$$

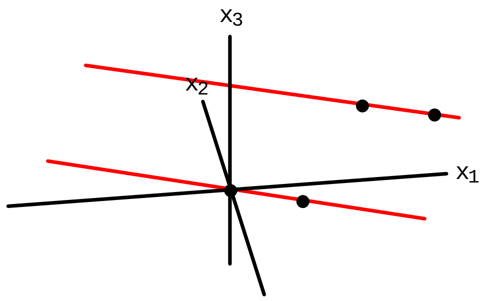
(same left side as in the previous example)

Solution:

$$\begin{bmatrix} 2 & 4 & -6 & 0 \\ 4 & 8 & -10 & 4 \end{bmatrix} \quad \text{row reduces to} \quad \begin{bmatrix} 1 & 2 & 0 & 6 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$

$$\mathbf{x} = \begin{bmatrix} 6 \\ 0 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \mathbf{p} + x_2 \mathbf{v}$$



Parallel solution sets of $A\mathbf{x} = \mathbf{0}$ & $A\mathbf{x} = \mathbf{b}$

6.1 Recap of Previous Two Examples

Solution of $A\mathbf{x} = \mathbf{0}$

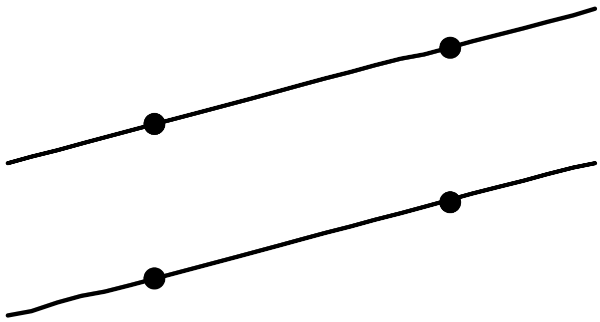
$$\mathbf{x} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = x_2\mathbf{v}$$

$\mathbf{x} = x_2\mathbf{v}$ = parametric equation of line passing through $\mathbf{0}$ and $\mathbf{v}$

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

$$\mathbf{x} = \begin{bmatrix} 6 \\ 0 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \mathbf{p} + x_2\mathbf{v}$$

$\mathbf{x} = \mathbf{p} + x_2\mathbf{v}$ = parametric equation of line passing through $\mathbf{p}$ parallel to $\mathbf{v}$



Parallel solution sets of
$A\mathbf{x} = \mathbf{b}$ and $A\mathbf{x} = \mathbf{0}$

Theorem 6.4

a) *Suppose* $\mathbf{p}$ *is a solution of* $A\mathbf{x} = \mathbf{b}$*, so that* $A\mathbf{p} = \mathbf{b}$*. Let* $\mathbf{v}_h$ *be any solution of the homogeneous equation* $A\mathbf{x} = \mathbf{0}$*, and let* $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$*. Then* $\mathbf{w}$ *is a solution of* $A\mathbf{x} = \mathbf{b}$*.*

b) *Every solution of* $A\mathbf{x} = \mathbf{b}$ *has the form* $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$*, with* $\mathbf{p}$ *a particular solution of* $A\mathbf{x} = \mathbf{b}$ *and* $\mathbf{v}_h$ *a solution of* $A\mathbf{x} = \mathbf{0}$*.*

Proof. Suppose $\mathbf{p}$ satisfies $A\mathbf{x} = \mathbf{b}$. Then $A\mathbf{p} = \mathbf{b}$. We claim that the solution set of $A\mathbf{x} = \mathbf{b}$ equals the set $S = \{\mathbf{w} : \mathbf{w} = \mathbf{p} + \mathbf{v}_h \quad \text{for some } \mathbf{v}_h \text{ such that } A\mathbf{v}_h = 0\ \}$. There are two things to prove:

a) every vector in S satisfies $A\mathbf{x} = \mathbf{b}$,

b) every vector that satisfies $A\mathbf{x} = \mathbf{b}$ is in S.

To prove *a*) let $\mathbf{w}$ have the form $\mathbf{w} = \mathbf{p}+\mathbf{v}_h$, where $A\mathbf{v}_h = 0$ and let $A\mathbf{p} = \mathbf{b}$. Then

$$A\mathbf{w} = A(\mathbf{p} + \mathbf{v_h}) = A\mathbf{p} + A\mathbf{v_h} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

So every vector of the form $\mathbf{p} + \mathbf{v}_h$, satisfies $A\mathbf{x} = \mathbf{b}$.

To prove *b*) let $\mathbf{w}$ be any solution of $A\mathbf{x} = \mathbf{b}$, and set $\mathbf{v}_h = \mathbf{w} - \mathbf{p}$. Then

$$A\mathbf{v}_h = A(\mathbf{w} - \mathbf{p}) = A\mathbf{w} - A\mathbf{p} = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

So $\mathbf{v}_h$ satisfies $A\mathbf{x} = \mathbf{0}$. Thus every solution of $A\mathbf{x} = \mathbf{b}$ has the form $\mathbf{w} = \mathbf{p}+\mathbf{v}_h$.
■

Example 6.5 *Describe the solution set of* $2x_1 - 4x_2 - 4x_3 = 0$; *compare it to the solution set* $2x_1 - 4x_2 - 4x_3 = 6$.

Solution: Corresponding augmented matrix to $2x_1 - 4x_2 - 4x_3 = 0$:

$$\begin{bmatrix} 2 & -4 & -4 & 0 \end{bmatrix} \sim \qquad \text{(fill-in)}$$

Vector form of the solution:

$$\mathbf{v} = \begin{bmatrix} 2x_2 + 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \_\_\_\_ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \_\_\_\_ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

Corresponding augmented matrix to $2x_1 - 4x_2 - 4x_3 = 6$:

$$\begin{bmatrix} 2 & -4 & -4 & 6 \end{bmatrix} \sim \qquad \text{(fill -in)}$$

Vector form of the solution (see Figure 6.1):

$$\mathbf{v} = \begin{bmatrix} 3 + 2x_2 + 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \end{bmatrix} + \_\_\_\_ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \_\_\_\_ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

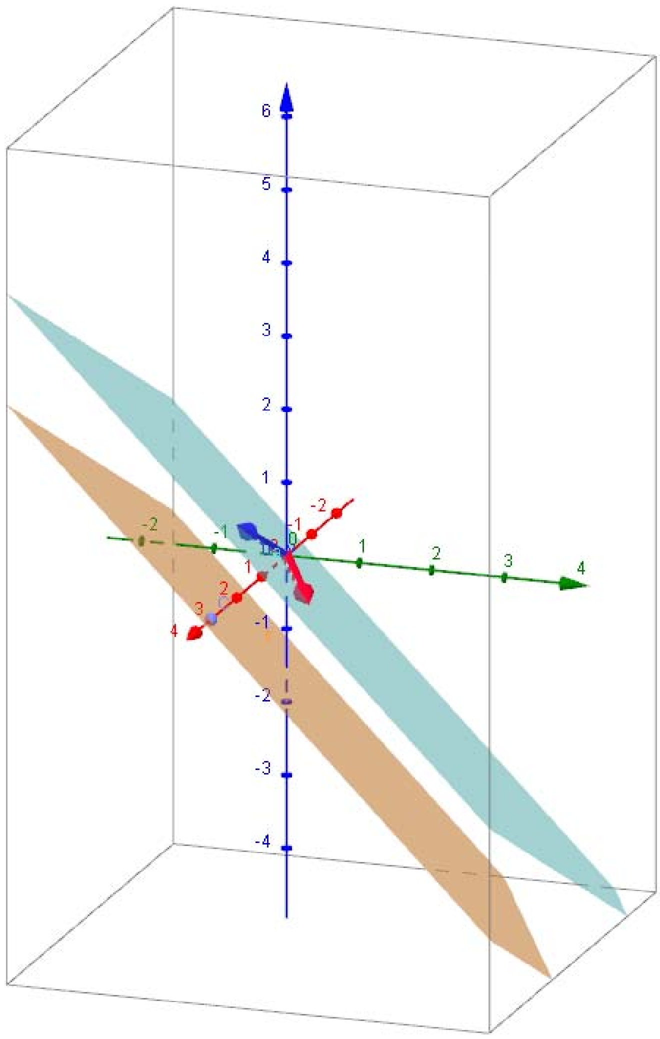


Fig. 6.1. Parallel Solution Sets of $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x} = \mathbf{b}$.

7

Linear Independence

The homogeneous equations in Chapter 6 can be studied from a different perspective by writing them as vector equations. In this way, the focus shifts from the unknown solutions of $A\mathbf{x} = \mathbf{0}$ to the vectors that appear in the vector equations.

A homogeneous system such as

$$\begin{bmatrix} 1 & 2 & -3 \\ 3 & 5 & 9 \\ 5 & 9 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

can be viewed as a vector equation

$$x_1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 9 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The vector equation has the trivial solution ($x_1 = 0$, $x_2 = 0$, $x_3 = 0$), but is this the *only solution*?

Definition 7.1 *A set of vectors* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ *in* $\mathbb{R}^n$ *is said to be* ***linearly independent*** *if the vector equation*

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution. The set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ *is said to be* ***linearly dependent*** *if there exists weights* $c_1, \ldots, c_p$*,not all* 0*, such that*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}.$$

$\uparrow$

linear dependence relation

(when weights are not all zero)

Example 7.2 *Let* $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} -3 \\ 9 \\ 3 \end{bmatrix}$

a) *Determine if* $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ *is linearly independent.*

b) *If possible, find a linear dependence relation among* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$*.*

Solution:

a)

$$x_1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 9 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Augmented matrix:

$$\begin{bmatrix} 1 & 2 & -3 & 0 \\ 3 & 5 & 9 & 0 \\ 5 & 9 & 3 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & -1 & 18 & 0 \\ 0 & -1 & 18 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & -1 & 18 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

x_3 is a free variable $\Rightarrow$ there are nontrivial solutions.

$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly dependent set

b) Reduced echelon form: $\begin{bmatrix} 1 & 0 & 33 & 0 \\ 0 & 1 & -18 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow \begin{array}{l} x_1 = \\ x_2 = \\ x_3 \end{array}$

Let $x_3 =$ _____ (any nonzero number). Then $x_1 =$ _____ and $x_2 =$ _____.

$$\_\_\_\_ \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + \_\_\_\_ \begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix} + \_\_\_\_ \begin{bmatrix} -3 \\ 9 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$\_\_\_\_\mathbf{v}_1 + \_\_\_\_\mathbf{v}_2 + \_\_\_\_\mathbf{v}_3 = \mathbf{0}$$

(one possible linear dependence relation)

□

7.1 Linear Independence of Matrix Columns

A linear dependence relation such as

$$-33\begin{bmatrix}1\\3\\5\end{bmatrix}+18\begin{bmatrix}2\\5\\9\end{bmatrix}+1\begin{bmatrix}-3\\9\\3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

can be written as the matrix equation:

$$\begin{bmatrix}1&2&-3\\3&5&9\\5&9&3\end{bmatrix}\begin{bmatrix}-33\\18\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Each linear dependence relation among the columns of A corresponds to a nontrivial solution to $A\mathbf{x}=\mathbf{0}$.

*The columns of matrix A are linearly independent if and only if the equation $A\mathbf{x}=\mathbf{0}$ has **only** the trivial solution.*

7.2 Special Cases

Sometimes we can determine linear independence of a set with minimal effort.

1. A Set of One Vector

Consider the set containing one nonzero vector: $\{\mathbf{v}_1\}$
The only solution to $x_1\mathbf{v}_1 = 0$ is $x_1 =$ _____.

So $\{\mathbf{v}_1\}$ is linearly independent when $\mathbf{v}_1 \neq \mathbf{0}$.

2. A Set of Two Vectors

Example 7.3 *Let*

$$\mathbf{u}_1=\begin{bmatrix}2\\1\end{bmatrix},\quad \mathbf{u}_2=\begin{bmatrix}4\\2\end{bmatrix},\quad \mathbf{v}_1=\begin{bmatrix}2\\1\end{bmatrix},\quad \mathbf{v}_2=\begin{bmatrix}2\\3\end{bmatrix}.$$

a) *Determine if* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *is a linearly dependent set or a linearly independent set.*

b) *Determine if* $\{\mathbf{v}_1, \mathbf{v}_2\}$ *is a linearly dependent set or a linearly independent set.*

Solution:

a) Notice that $\mathbf{u}_2 =$ _____ $\mathbf{u}_1$. Therefore

_____ $\mathbf{u}_1 +$ _____ $\mathbf{u}_2 = 0$

This means that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a linearly __________________ set.

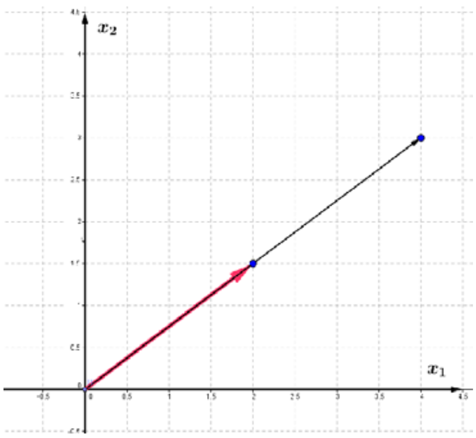
b) Suppose

$$c\mathbf{v}_1 + d\mathbf{v}_2 = \mathbf{0}.$$

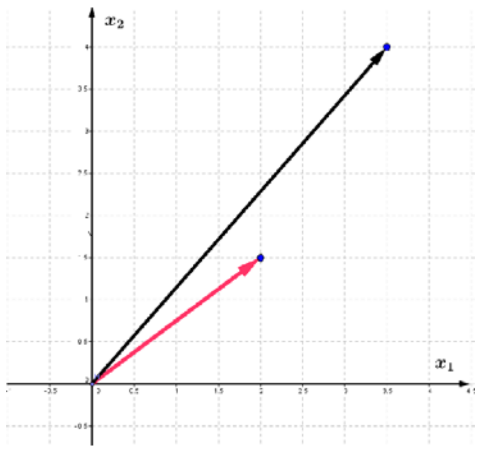
Then $\mathbf{v}_1 =$ ——— $\mathbf{v}_2$ if $c \neq 0$. But this is impossible since $\mathbf{v}_1$ is - ______ a multiple of $\mathbf{v}_2$ which means $c =$ _____.

Similarly, $\mathbf{v}_2 =$ ——— $\mathbf{v}_1$ if $d \neq 0$. But this is impossible since $\mathbf{v}_2$ is not a multiple of $\mathbf{v}_1$ and so $d = 0$. This means that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly __________________ set.

A set of two vectors is linearly dependent if at least one vector is a multiple of the other.
A set of two vectors is linearly independent if and only if neither of the vectors is a multiple of the other.



linearly __________________



linearly __________________

3. A Set Containing the $\mathbf{0}$ Vector

Theorem 7.4 *A set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ containing the zero vector is linearly dependent.*

Proof: Renumber the vectors so that $\mathbf{v}_1 =$ ____. Then

$$\_\_\_\_\mathbf{v}_1 + \_\_\_\_\_\mathbf{v}_2 + \cdots + \_\_\_\_\_\mathbf{v}_p = \mathbf{0}$$

which shows that S is linearly _________________.

4. A Set Containing Too Many Vectors

Theorem 7.5 *If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. I.e. any set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.*

Outline of Proof:

$$A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_p \end{bmatrix} \text{ is } n \times p$$

Suppose $p > n$.

$$\Longrightarrow A\mathbf{x} = \mathbf{0} \text{ has more variables than equations}$$
$$\Longrightarrow A\mathbf{x} = \mathbf{0} \text{ has nontrivial solutions}$$
$$\Longrightarrow \text{columns of } A \text{ are linearly dependent}$$

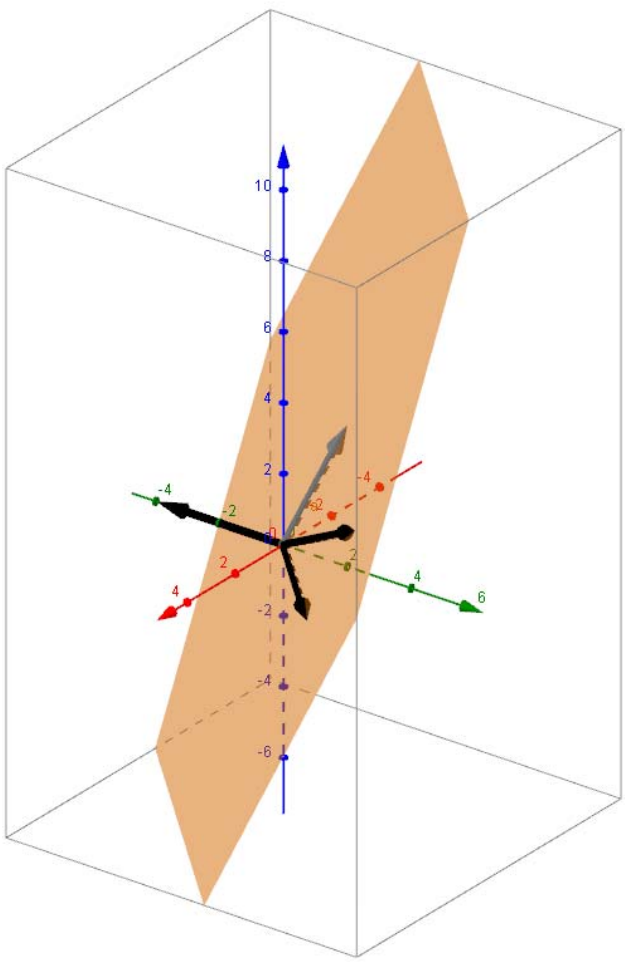
Example 7.6 *With the least amount of work possible, decide which of the following sets of vectors are linearly independent and give a reason for each answer.*

a. $\left\{ \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 9 \\ 6 \\ 4 \end{bmatrix} \right\}$ b. Columns of $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 0 \\ 9 & 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 & 8 \end{bmatrix}$

c. $\left\{ \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 9 \\ 6 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$ d. $\left\{ \begin{bmatrix} 8 \\ 2 \\ 1 \\ 4 \end{bmatrix} \right\}$

7.3 Characterization of Linearly Dependent Sets

Example 7.7 *Consider the set of vectors* $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ *in* $\mathbb{R}^3$ *in the following diagram (all in black). Is the set linearly dependent? Explain.*



Theorem 7.8 *An indexed set* $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ *of two or more vectors is linearly dependent if and only if at least one of the vectors in* S *is a linear combination of the others. In fact, if* S *is linearly dependent, and* $\mathbf{v}_1 \neq \mathbf{0}$, *then some vector* $\mathbf{v}_j$ *(*$j \geq 2$*) is a linear combination of the preceding vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$.

8

Introduction to Linear Transformations

In this chapter we shall begin the study of functions of the form $\mathbf{w} = F(\mathbf{x})$, where the independent variable $\mathbf{x}$ is a vector in $\mathbb{R}^n$ and the dependent variable $\mathbf{w}$ is a vector in $\mathbb{R}^m$. We shall concentrate on a special class of such functions called "*linear transformations*." Linear transformations are fundamental in the study of linear algebra and have many important applications in physics, engineering, social sciences, and various branches of mathematics.

The difference between a matrix equation $A\mathbf{x} = \mathbf{b}$ and the associated vector equation

$$x_1\mathbf{a}_1 + ... + x_n\mathbf{a}_n = \mathbf{b}$$

is merely a matter of notation. However, a matrix equation $A\mathbf{x} = \mathbf{b}$ can arise in linear algebra (and in applications such as computer graphics and signal processing) in a way that is not directly connected with linear combinations of vectors. This happens when we think of the matrix A as an object that "acts" on a vector $\mathbf{x}$ by multiplication to produce a new vector called $A\mathbf{x}$.

<<<< * >>>>

Another way to view $A\mathbf{x} = \mathbf{b}$:

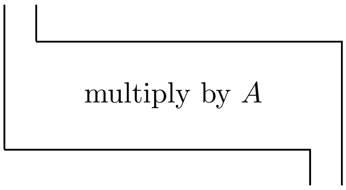
Matrix A is an object acting on $\mathbf{x}$ by multiplication to produce a new vector $A\mathbf{x}$ or $\mathbf{b}$.

Example 8.1

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

□

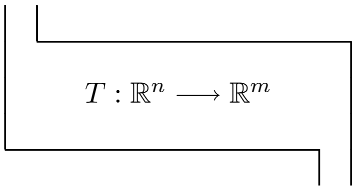
Suppose A is $m \times n$. Solving $A\mathbf{x} = \mathbf{b}$ amounts to finding all ____ in $\mathbb{R}^n$ which are transformed into vector $\mathbf{b}$ in $\mathbb{R}^m$ through multiplication by A.



transformation "machine"

8.1 Matrix Transformations

A **transformation** T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$.



Terminology:

$\mathbb{R}^n$: **domain** of T $\qquad$ $\mathbb{R}^m$: **codomain** of T

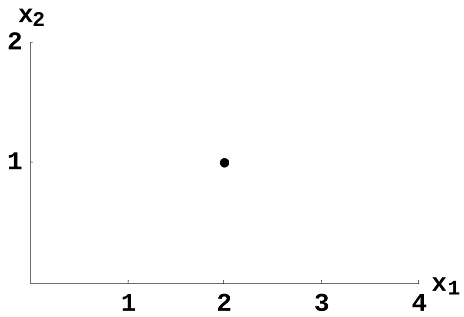
$T(\mathbf{x})$ in $\mathbb{R}^m$ is the **image** of $\mathbf{x}$ under the transformation T

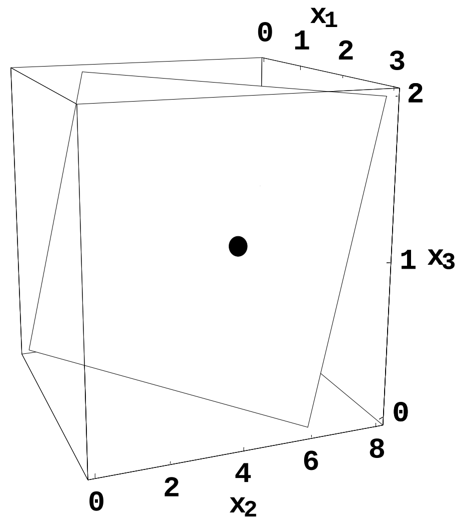
Set of all images $T(\mathbf{x})$ is the **range** of T

Example 8.2 *Let* $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$. *Define a transformation* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *by* $T(\mathbf{x}) = A\mathbf{x}$.

Then if $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$,

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}$$





Exercise 8.1 *Let* $A = \begin{bmatrix} 1 & -2 & 3 \\ -5 & 10 & -15 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ -10 \end{bmatrix}$ *and* $\mathbf{c} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$. *Then define a transformation* $T : \mathbb{R}^3 \to \mathbb{R}^2$ *by* $T(\mathbf{x}) = A\mathbf{x}$.

a) *Find an* $\mathbf{x}$ *in* $\mathbb{R}^3$ *whose image under* T *is* $\mathbf{b}$.

b) *Is there more than one* $\mathbf{x}$ *under* T *whose image is* $\mathbf{b}$. *(uniqueness problem)*

c) *Determine if* $\mathbf{c}$ *is in the range of the transformation* T. *(existence problem)*

Solution:

a) Solve _______=_____ for $\mathbf{x}$. I.e., solve _______=_____ or

$$\begin{bmatrix} 1 & -2 & 3 \\ -5 & 10 & -15 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -10 \end{bmatrix}$$

Augmented matrix:

$$\begin{bmatrix} 1 & -2 & 3 & 2 \\ -5 & 10 & -15 & -10 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$x_1 = 2x_2 - 3x_3 + 2$
x_2 is free
x_3 is free

Let $x_2 =$ _____ and $x_3 =$ _____. Then $x_1 =$ _____. So

$$\mathbf{x} = \begin{bmatrix} \\ \\ \end{bmatrix}$$

b) Is there an $\mathbf{x}$ for which $T(\mathbf{x}) = \mathbf{b}$?

Free variables exist $\Longrightarrow$ There is more than one $\mathbf{x}$ for which $T(\mathbf{x}) = \mathbf{b}$

c) Is there an $\mathbf{x}$ for which $T(\mathbf{x}) = \mathbf{c}$? This is another way of asking if $A\mathbf{x} = \mathbf{c}$ is ________________.
Augmented matrix:

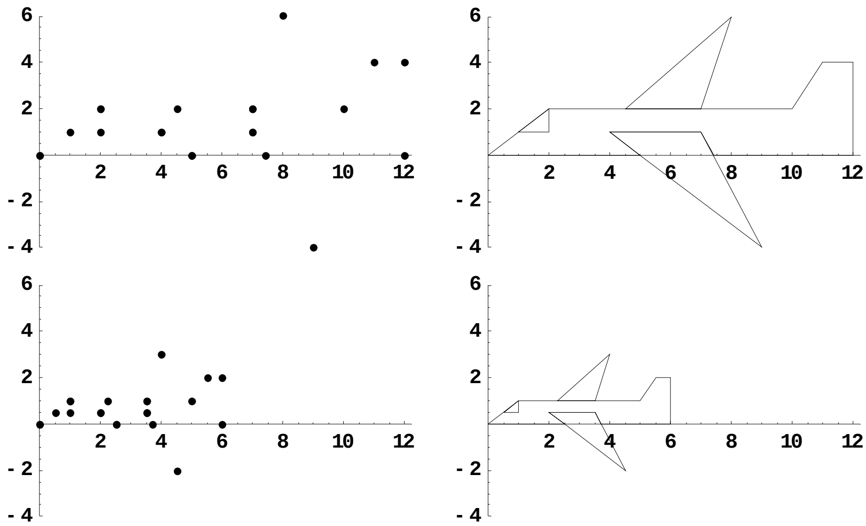
$$\begin{bmatrix} 1 & -2 & 3 & 3 \\ -5 & 10 & -15 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$\mathbf{c}$ is not in the ________________ of T. □

Matrix transformations have many applications - including *computer graphics*.

Example 8.3 *Let* $A = \begin{bmatrix} .5 & 0 \\ 0 & .5 \end{bmatrix}$. *The transformation* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *defined by* $T(\mathbf{x}) = A\mathbf{x}$ *is an example of a* ***contraction*** *transformation. The transformation* $T(\mathbf{x}) = A\mathbf{x}$ *can be used to move a point* $\mathbf{x}$.

$$\mathbf{u} = \begin{bmatrix} 8 \\ 6 \end{bmatrix} \qquad T(\mathbf{u}) = \begin{bmatrix} .5 & 0 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} 8 \\ 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$



8.2 Linear Transformations

If A is $m \times n$, then the transformation $T(\mathbf{x}) = A\mathbf{x}$ has the following properties:

$$T(\mathbf{u}+\mathbf{v}) = A(\mathbf{u}+\mathbf{v}) = \_\_\_\_\_\_\_ + \_\_\_\_\_\_\_$$

$$= \_\_\_\_\_\_ + \_\_\_\_\_\_$$

and

$$T(c\mathbf{u}) = A(c\mathbf{u}) = \_\_\_\_\_ A\mathbf{u} = \_\_\_\_\_ T(\mathbf{u})$$

for all $\mathbf{u}$,$\mathbf{v}$ in $\mathbb{R}^n$ and all scalars c.

Definition 8.4 *A transformation T is* ***linear*** *if:*

i. $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T.

ii. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all $\mathbf{u}$ in the domain of T and all scalars c.

Every matrix transformation is a **linear** transformation.

Theorem 8.5 *If T is a linear transformation, then*

$$T(\mathbf{0}) = \mathbf{0} \qquad \textit{and} \qquad T(c\mathbf{u}+d\mathbf{v}) = c\mathbf{T}(\mathbf{u}) + d\mathbf{T}(\mathbf{v}).$$

Proof.

$$T(\mathbf{0}) = T(0\mathbf{u}) = \_\_\_\_T(\mathbf{u}) = \_\_\_\_\_.$$

$$T(c\mathbf{u} + d\mathbf{v}) = T(\qquad) + T(\qquad) = \_\_\_\_\_T(\qquad) + \_\_\_\_\_T(\qquad)$$

■

Example 8.6 *Let* $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{y}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ *and* $\mathbf{y}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$. *Suppose* $T : \mathbb{R}^2 \to \mathbb{R}^3$ *is a linear transformation which maps* $\mathbf{e}_1$ *into* $\mathbf{y}_1$ *and* $\mathbf{e}_2$ *into* $\mathbf{y}_2$. *Find the images of* $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ *and* $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Solution: First, note that

$$T(\mathbf{e}_1) = \_\_\_\_\_\_ \qquad \text{and} \qquad T(\mathbf{e}_2) = \_\_\_\_\_\_.$$

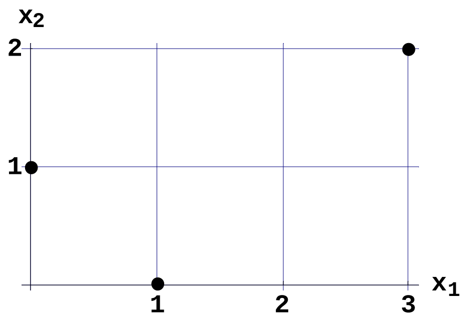
Also

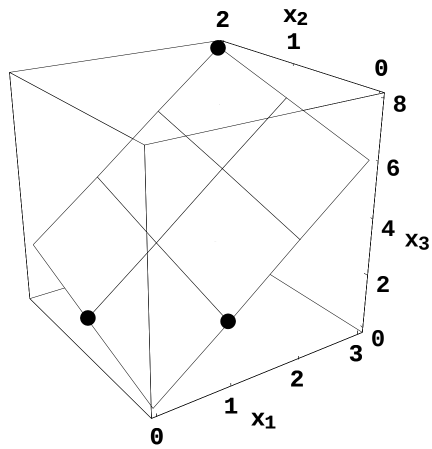
$$\_\_\_\mathbf{e}_1 + \_\_\_\mathbf{e}_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Then

$$T\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}\right) = T(\_\_\_\mathbf{e}_1 + \_\_\_\mathbf{e}_2) = \_\_\_T(\mathbf{e}_1) + \_\_\_T(\mathbf{e}_2)$$

$$=$$





$$T(3\mathbf{e}_1 + 2\mathbf{e}_2) = 3T(\mathbf{e}_1) + 2T(\mathbf{e}_2)$$

Also

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = T(\_\_\_\_\_\_\mathbf{e}_1 + \_\_\_\_\_\_\mathbf{e}_2) =$$
$$\_\_\_\_\_T(\mathbf{e}_1) + \_\_\_\_\_T(\mathbf{e}_2) =$$

$\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$ □

Example 8.7 *Define $T : \mathbb{R}^3 \to \mathbb{R}^2$ such that*

$$T(x_1, x_2, x_3) = (|x_1 + x_3|,\ 2 + 5x_2).$$

Show that T is a not a linear transformation.

Solution: Another way to write the transformation:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} |x_1 + x_3| \\ 2 + 5x_2 \end{bmatrix}$$

Provide a **counterexample** - example where $T(\mathbf{0}) = \mathbf{0}$, $T(c\mathbf{u}) = c\mathbf{T}(\mathbf{u})$ or $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ is violated.

A counterexample:

$$T(\mathbf{0}) = T\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \quad\quad\quad \end{bmatrix} \neq \_\_\_\_\_\_$$

which means that T is not linear.

Another counterexample: Let $c = -1$ and $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Then

$$T(c\mathbf{u}) = T\left(\begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} |-1+-1| \\ 2+5(-1) \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

and

$$cT(\mathbf{u}) = -1T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = -1\begin{bmatrix} \quad \end{bmatrix} = \begin{bmatrix} \quad \end{bmatrix}.$$

Therefore $T(c\mathbf{u}) \neq$ ___ $T(\mathbf{u})$ and therefore T is not ______________.

□

9

The Matrix of a Linear Transformation

Identity Matrix I_n is an $n \times n$ matrix with 1's on the main left to right diagonal and 0's elsewhere. The ith column of I_n is labeled $\mathbf{e}_i$.

Example 9.1

$$I_3 = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

□

Note that

$$I_3\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \_\_\_\_ \begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix} + \_\_\_\_ \begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix} + \_\_\_\_ \begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix} = \_\_\_\_.$$

In general, for $\mathbf{x}$ in $\mathbb{R}^n$,

$$I_n\mathbf{x} = \_\_\_$$

From chapter 7, if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then

$$T(c\mathbf{u} + d\mathbf{v}) = c\mathbf{T}(\mathbf{u}) + d\mathbf{T}(\mathbf{v}).$$

Generalized Result:

$$T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p).$$

Example 9.2 *The columns of* $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ *are* $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *and* $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. *Suppose* T *is a linear transformation from* $\mathbb{R}^2$ *to* $\mathbb{R}^3$ *where*

$$T(\mathbf{e}_1) = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix}.$$

Compute $T(\mathbf{x})$ *for any* $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Solution: A vector $\mathbf{x}$ in $\mathbb{R}^2$ can be written as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \_\_\_\_\_ \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \_\_\_\_\_ \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \_\_\_\_\_\mathbf{e}_1 + \_\_\_\_\_\mathbf{e}_2$$

Then

$$T(\mathbf{x}) = T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = \_\_\_\_\_T(\mathbf{e}_1) + \_\_\_\_\_T(\mathbf{e}_2)$$

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

Note that

$$T(\mathbf{x}) = \begin{bmatrix} \qquad \\ \\ \\ \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

So

$$T(\mathbf{x}) = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} \mathbf{x} = A\mathbf{x}$$

To get A*, replace the identity matrix* $\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 \end{bmatrix}$ *with* $\begin{bmatrix} T(\mathbf{e}_2) & T(\mathbf{e}_2) \end{bmatrix}$.

Theorem 9.3 *Let* $T : \mathbb{R}^n \to \mathbb{R}^m$ *be a linear transformation. Then there exists a unique matrix* A *such that*

$$T(\mathbf{x}) = A\mathbf{x} \quad \textit{for all} \quad \mathbf{x} \textit{ in } \mathbb{R}^n.$$

In fact, A *is the* $m \times n$ *matrix whose* j*-th column is the vector* $T(\mathbf{e}_j)$*, where* $\mathbf{e}_j$ *is the* j*-th column of the identity matrix in* $\mathbb{R}^n$*.*

$$A = [T(\mathbf{e}_1) \qquad T(\mathbf{e}_2) \qquad \cdots \qquad T(\mathbf{e}_n)]$$

$$\uparrow$$

standard matrix for the linear transformation T

$$\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 2x_2 \\ 4x_1 \\ 3x_1 + 2x_2 \end{bmatrix}$$

Solution:

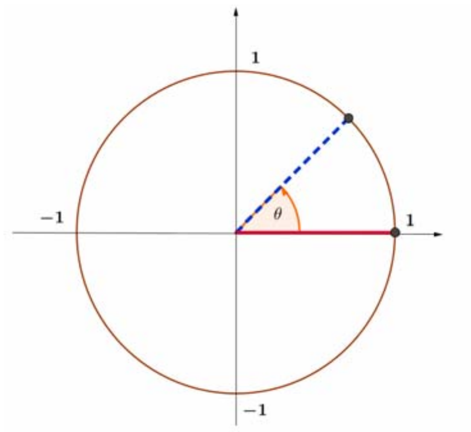
$$\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix} = \text{standard matrix of the linear transformation } T$$

$$\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix} = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} =$$

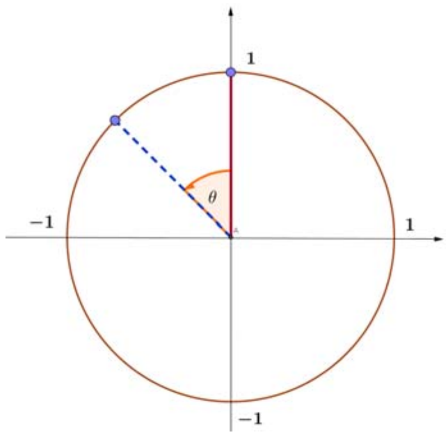
(fill-in)

□

Example 9.4 *Find the standard matrix of the linear transformation* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *which rotates a point about the origin through an angle of* $\frac{\pi}{4}$ *radians (counterclockwise).*



$$T(\mathbf{e}_1) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$



$$T(\mathbf{e}_2) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

$$\Longrightarrow \qquad A = \begin{bmatrix} \qquad\qquad\qquad \end{bmatrix}$$

10
Matrix Operations

10.1 Matrix Notation

Two ways to denote $m \times n$ matrix A:

a) In terms of the *columns* of A:

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$$

b) In terms of the *entries* of A:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

Main diagonal entries:____________________

Zero matrix:

$$0 = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

Theorem 10.1 *Let A, B, and C be matrices of the same size, and let r and s be scalars. Then*

a)	$A + B = B + A$	*d)*	$r(A + B) = rA + rB$
b)	$(A + B) + C = A + (B + C)$	*e)*	$(r + s)\,A = rA + sA$
c)	$A + 0 = A$	*f)*	$r\,(sA) = (rs)\,A$

10.2 Matrix Multiplication

Multiplying B and $\mathbf{x}$ transforms $\mathbf{x}$ into the vector $B\mathbf{x}$. In turn, if we multiply A and $B\mathbf{x}$, we transform $B\mathbf{x}$ into $A\,(B\mathbf{x})$. So $A\,(B\mathbf{x})$ is the composition of two mappings.

Define the product AB so that $A\,(B\mathbf{x}) = (AB)\,\mathbf{x}$.

Suppose A is $m \times n$ and B is $n \times p$ where

$$B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_p] \text{ and } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}.$$

Then

$$B\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_p\mathbf{b}_p$$

and

$$A\,(B\mathbf{x}) = A(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_p\mathbf{b}_p)$$

$$= A\,(x_1\mathbf{b}_1) + A\,(x_2\mathbf{b}_2) + \cdots + A\,(x_p\mathbf{b}_p)$$

$$= x_1A\mathbf{b}_1 + x_2A\mathbf{b}_2 + \cdots + x_pA\mathbf{b}_p = [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}.$$

Therefore,

$$A\,(B\mathbf{x}) = [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]\,\mathbf{x}.$$

and by defining

$$\boxed{AB = [A\mathbf{b}_1 \ \ A\mathbf{b}_2 \ \ \cdots \ \ A\mathbf{b}_p]}$$

we have $A(B\mathbf{x}) = (AB)\mathbf{x}$.

Example 10.2 *Compute AB where $A = \begin{bmatrix} 4 & -2 \\ 3 & -5 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -3 \\ 6 & -7 \end{bmatrix}$.*

Solution:

$$A\mathbf{b}_1 = \begin{bmatrix} 4 & -2 \\ 3 & -5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \end{bmatrix} = \begin{bmatrix} -4 \\ -24 \\ 6 \end{bmatrix}, \qquad A\mathbf{b}_2 = \begin{bmatrix} 4 & -2 \\ 3 & -5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -7 \end{bmatrix} = \begin{bmatrix} 2 \\ 26 \\ -7 \end{bmatrix}$$

$$\Longrightarrow AB = \begin{bmatrix} -4 & 2 \\ -24 & 26 \\ 6 & -7 \end{bmatrix}$$

Note that $A\mathbf{b}_1$ is a linear combination of the columns of A and $A\mathbf{b}_2$ is a linear combination of the columns of A.

> Each column of AB is a linear combination of the columns of A using weights from the corresponding columns of B.

Example 10.3 *If A is 4×3 and B is 3×2, then what are the sizes of AB and BA?*

Solution:

$$AB = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix} = \begin{bmatrix} & \\ & \\ & \\ & \end{bmatrix}$$

BA would be $\begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix} \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$ which is ___________________.

$$\boxed{\text{If } A \text{ is } m \times n \text{ and } B \text{ is } n \times p, \text{ then } AB \text{ is } m \times p.}$$

10.3 Row-Column Rule for Computing AB (alternate method)

The definition

$$\boxed{AB = [A\mathbf{b}_1 \;\; A\mathbf{b}_2 \;\; \cdots \;\; A\mathbf{b}_p]}$$

is good for theoretical work.

When A and B have small sizes, the following method is more efficient when working by hand.

If AB is defined, let $(AB)_{ij}$ denote the entry in the ith row and jth column of AB. Then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

$$\begin{bmatrix} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \\ \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = \begin{bmatrix} \\ (AB)_{ij} \\ \\ \end{bmatrix}$$

Example 10.4 *Let* $A = \begin{bmatrix} 2 & 3 & 6 \\ -1 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & -3 \\ 0 & 1 \\ 4 & -7 \end{bmatrix}$. *Compute* AB*, if it is defined.*

Solution: Since A is 2×3 and B is 3×2, then AB is defined and AB is _____$\times$_____.

$AB = \begin{bmatrix} \mathbf{2} & \mathbf{3} & \mathbf{6} \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{2} & -3 \\ \mathbf{0} & 1 \\ \mathbf{4} & -7 \end{bmatrix} = \begin{bmatrix} \mathbf{28} & \blacksquare \\ \blacksquare & \blacksquare \end{bmatrix}$, $\begin{bmatrix} \mathbf{2} & \mathbf{3} & \mathbf{6} \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & \mathbf{-3} \\ 0 & \mathbf{1} \\ 4 & \mathbf{-7} \end{bmatrix} = \begin{bmatrix} 28 & \mathbf{-45} \\ \blacksquare & \blacksquare \end{bmatrix}$

$\begin{bmatrix} 2 & 3 & 6 \\ \mathbf{-1} & \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{2} & -3 \\ \mathbf{0} & 1 \\ \mathbf{4} & -7 \end{bmatrix} = \begin{bmatrix} 28 & -45 \\ \mathbf{2} & \blacksquare \end{bmatrix}$, $\begin{bmatrix} 2 & 3 & 6 \\ \mathbf{-1} & \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 2 & \mathbf{-3} \\ 0 & \mathbf{1} \\ 4 & \mathbf{-7} \end{bmatrix} = \begin{bmatrix} 28 & -45 \\ 2 & \mathbf{-4} \end{bmatrix}$

So $AB = \begin{bmatrix} 28 & -45 \\ 2 & -4 \end{bmatrix}$. $\square$

Theorem 10.5 *Let A be $m \times n$ and let B and C have sizes for which the indicated sums and products are defined.*

a)	$A(BC) = (AB)C$	(associative law of multiplication)
b)	$A(B + C) = AB + AC$	(left - distributive law)
c)	$(B + C)A = BA + CA$	(right-distributive law)
d)	$r(AB) = (rA)B = A(rB)$ for any scalar r	
e)	$I_m A = A = AI_n$	(identity for matrix multiplication)

WARNINGS:

Properties above are analogous to properties of real numbers. But **NOT ALL** real number properties correspond to matrix properties.

1. It is not the case that AB always equal BA. (see Exercises)
2. Even if $AB = AC$, then B may not equal C. (see Exercises)
3. It is possible for $AB = 0$ even if $A \neq 0$ and $B \neq 0$. (see Exercises)

10.4 Powers of A

$$A^k = \underbrace{A \cdots A}_{k}$$

$$\begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}^3 = \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 21 & 8 \end{bmatrix}$$

If A is $m \times n$, the **transpose** of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 8 \\ 7 & 6 & 5 & 4 & 3 \end{bmatrix} \quad \Longrightarrow \quad A^T = \begin{bmatrix} 1 & 6 & 7 \\ 2 & 7 & 6 \\ 3 & 8 & 5 \\ 4 & 9 & 4 \\ 5 & 8 & 3 \end{bmatrix}$$

Example 10.6 *Let* $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -2 & 4 \end{bmatrix}$. *Compute* AB, $(AB)^T$, $A^T B^T$ *and* $B^T A^T$.

Solution:

$$AB = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} \quad & \quad & \quad \end{bmatrix}$$

$$(AB)^T = \begin{bmatrix} \quad & \quad & \quad \end{bmatrix}$$

$$A^T B^T = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 3 & 10 \\ 2 & 0 & -4 \\ 2 & 1 & 4 \end{bmatrix}$$

$$B^T A^T = \begin{bmatrix} 1 & 0 & -2 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \quad & \quad & \quad \end{bmatrix}$$

Theorem 10.7 *Let A and B denote matrices whose sizes are appropriate for the following sums and products. Then*

a) $\left(A^T\right)^T = A$ *(i.e., the transpose of A^T is A)*

b) $(A + B)^T = A^T + B^T$

c) *For any scalar r, $(rA)^T = rA^T$*

d) $(AB)^T = B^T A^T$ *(I.e. the transpose of a product of matrices equals the product of their transposes in reverse order.)*

Example 10.8 *Prove that $(ABC)^T =$ __________.*

Solution: By Theorem 10.7,

$$(ABC)^T = ((AB)\,C)^T = C^T \left(\qquad \right)^T = C^T \left(\qquad \right) =$$

__________. □

11

The Inverse of a Matrix

11.1 Invertible Matrices

The inverse of a real number a is denoted by a^{-1}. For example, $7^{-1} = 1/7$ and

$$7 \cdot 7^{-1} = 7^{-1} \cdot 7 = 1$$

An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C satisfying

$$CA = AC = I_n$$

where I_n is the $n \times n$ identity matrix. We call C the **inverse** of A .

FACT If A is invertible, then the inverse is unique.

Proof. *Proof:* Assume B and C are both inverses of A. Then

$$B = BI = B\,(\_\_\_\_\_) = (\_\_\_\_\_\_)\,\_\_\_\_\_\_ = I\_\_\_\_\_\_ = C.$$

So the inverse is unique since any two inverses coincide. ■

The inverse of A is usually denoted by A^{-1}.

We have

$$\boxed{AA^{-1} = A^{-1}A = I_n}$$

Not all $n \times n$ matrices are invertible. A matrix which is *not* invertible is sometimes called a **singular** matrix. An invertible matrix is called **nonsingular** matrix.

Theorem 11.1 *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *If* $ad - bc \neq 0$, *then* A *is invertible and*

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

If $ad - bc = 0$, *then* A *is not invertible.*

Assume A is any invertible matrix and we wish to solve $A\mathbf{x} = \mathbf{b}$. Then

$$\_\_\_\_\_A\mathbf{x} = \_\_\_\_\_\mathbf{b} \qquad \text{and so}$$

$$I\mathbf{x} = \_\_\_\_\_\_\_\_ \text{ or } \mathbf{x} = \_\_\_\_\_\_\_\_.$$

Suppose $\mathbf{w}$ is also a solution to $A\mathbf{x} = \mathbf{b}$. Then $A\mathbf{w} = \mathbf{b}$ and

$$\_\_\_\_\_A\mathbf{w} = \_\_\_\_\_\mathbf{b} \qquad \text{which means} \qquad \mathbf{w} = A^{-1}\mathbf{b}.$$

So, $\mathbf{w} = A^{-1}\mathbf{b}$, which is in fact the same solution.
We have proved the following result:

Theorem 11.2 *If A is an invertible $n \times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.*

Proof. Take any $\mathbf{b}$ in $\mathbb{R}^n$. A solution exists because if $A^{-1}\mathbf{b}$ is substituted for $\mathbf{x}$, then $A\mathbf{x} = A\left(A^{-1}\mathbf{b}\right) = \left(AA^{-1}\right)\mathbf{b} = I\mathbf{b} = \mathbf{b}$. So $A^{-1}\mathbf{b}$ is a solution. To prove that the solution is unique, show that if $\mathbf{u}$ is any solution, then $\mathbf{u}$ in fact, must be $A^{-1}\mathbf{b}$. Indeed, if $A\mathbf{u} = b$, we can multiply both sides by A^{-1} and obtain

$$A^{-1}A\mathbf{u} = A^{-1}\mathbf{b}, \quad I\mathbf{u} = A^{-1}\mathbf{b}, \quad \mathbf{u} = A^{-1}\mathbf{b}.$$

■

Example 11.3 *Use the inverse of* $A = \begin{bmatrix} -7 & 3 \\ 5 & -2 \end{bmatrix}$ *to solve*

$$\begin{array}{rcrcr} -7x_1 & + & 3x_2 & = & 2 \\ 5x_1 & - & 2x_2 & = & 1 \end{array}.$$

Solution: Matrix form of the linear system: $\begin{bmatrix} -7 & 3 \\ 5 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$$A^{-1} = \tfrac{1}{14-15} \begin{bmatrix} -2 & -3 \\ -5 & -7 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}.$$

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} \quad \\ \quad \end{bmatrix} = \begin{bmatrix} \quad \\ \quad \end{bmatrix}.$$

Theorem 11.4 *Suppose A and B are invertible. Then the following results hold:*

a) A^{-1} *is invertible and* $\left(A^{-1}\right)^{-1} = A$ *(i.e. A is the inverse of A^{-1}).*

b) AB *is invertible and* $(AB)^{-1} = B^{-1}A^{-1}$

c) A^T *is invertible and* $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

Partial proof of part b:

$$(AB)\left(B^{-1}A^{-1}\right) = A\,(\_\_\_\_\_\_\_\_\_\_)\,A^{-1}$$

$$= A\,(\_\_\_\_\_\_\_\_\_\_\_)\,A^{-1} = \_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_.$$

Similarly, one can show that $\left(B^{-1}A^{-1}\right)(AB) = I$.

Theorem 11.4, part b) can be generalized to three or more invertible matrices:

$$(ABC)^{-1} = \_\_\_\_\_\_\_\_\_\_\_$$

Earlier, we saw a formula for finding the inverse of a 2×2 invertible matrix. How do we find the inverse of an invertible $n \times n$ matrix? To answer this question, we first look at **elementary** matrices.

11.2 Elementary Matrices

Definition 11.5 *An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix.*

Example 11.6 *Let*

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

$$\text{and} \quad A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}.$$

Then E_1, E_2, and E_3 are elementary matrices. Why? Observe the following products and describe how these products can be obtained by elementary row *operations on A.*

$$E_1A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ 2d & 2e & 2f \\ g & h & i \end{bmatrix}$$

$$E_2A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ g & h & i \\ d & e & f \end{bmatrix}$$

$$E_3A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 3a+g & 3b+h & 3c+i \end{bmatrix}$$

If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operations on I_m.

Elementary matrices are *invertible* because row operations are *reversible.* To determine the inverse of an elementary matrix E, determine the elementary row operation needed to transform E back into I and apply this operation to I to find the inverse.

For example,

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \qquad E_3^{-1} = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

Example 11.7 *Let* $A = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}$. *Then*

$$E_1A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$E_2\left(E_1A\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -3 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$E_3\left(E_2E_1A\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

So

$$\boxed{E_3E_2E_1A = I_3}.$$

Then multiplying on the right by A^{-1}, we get

$$E_3E_2E_1A\_\_\_\_\_\_ = I_3\_\_\_\_\_\_.$$

So

$$\boxed{E_3E_2E_1I_3 = A^{-1}}$$

The elementary row operations that row reduce A *to* $\mathbf{I}_n$ *are the same elementary row operations that transform* $\mathbf{I}_n$ *into* $\mathbf{A}^{-1}$.

Theorem 11.8 *An* $n \times n$ *matrix* A *is invertible if and only if* A *is row equivalent to* I_n*, and in this case, any sequence of elementary row operations that reduces* A *to* I_n *will also transform* I_n *to* A^{-1}.

11.3 Another View of Matrix Inversion

Denote the columns of I_n by $\mathbf{e}_1, ..., \mathbf{e}_n$. Then row reduction of $[A \;\; I]$ to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$ can be viewed as the simultaneous solution of the n systems

$$A\mathbf{x} = \mathbf{e}_1,\; A\mathbf{x} = \mathbf{e}_2, ..., A\mathbf{x} = \mathbf{e}_n, \tag{11.1}$$

where the "augmented columns" of these systems have all been placed next to A to form

$$\begin{bmatrix} A & \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_n \end{bmatrix} = [A \;\; I].$$

The equation $AA^{-1} = I$ and the definition of matrix multiplication show that the columns of A^{-1} are precisely the solutions of the systems in (11.1). This observation is useful because some applied problems may require finding only one or two columns of A^{-1}. In this case, only the corresponding systems in (11.1) need be solved.

11.4 Algorithm for finding A^{-1}

Place A and I side-by-side to form an augmented matrix $[A \;\; I]$. Then perform row operations on this matrix (which will produce identical operations on A and I). So by Theorem 11.8:

$$[A \;\; I] \text{ will row reduce to } \begin{bmatrix} I & A^{-1} \end{bmatrix}$$

or A is not invertible.

Example 11.9 *Find the inverse of*

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

if it exists.

Solution:

$$[A \;\; I] = \begin{bmatrix} 2 & 0 & 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{3}{2} & 1 & 0 \end{bmatrix}$$

So $A^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \\ \frac{3}{2} & 1 & 0 \end{bmatrix}$

Order of multiplication is important!

Example 11.10 *Suppose A,B,C, and D are invertible $n \times n$ matrices and $A = B(D - I_n)C$. Solve for D in terms of A, B, C and D.*

Solution:

$$\_\_\_\_\_A\_\_\_\_\_ = \_\_\_\_\_B(D - I_n)C\_\_\_\_\_$$

$$D - I_n = B^{-1}AC^{-1}$$

$$D - I_n + \_\_\_\_ = B^{-1}AC^{-1} + \_\_\_\_$$

$$D = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

12
Characterizations of Invertible Matrices

12.1 The Invertible Matrix Theorem

Theorem 12.1 ***(The Invertible Matrix Theorem)*** *Let A be a square $n \times n$ matrix. The the following statements are equivalent (i.e., for a given A, they are either all true or all false).*

a. A is an invertible matrix.

b. A is row equivalent to I_n.

c. A has n pivot positions.

d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

e. The columns of A form a linearly independent set.

f. The linear transformation $\mathbf{x} \rightarrow A\mathbf{x}$ is one-to-one.

g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

h. The columns of A span $\mathbb{R}^n$.

i. The linear transformation $\mathbf{x} \rightarrow A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

j. There is an $n \times n$ matrix C such that $CA = I_n$.

k. There is an $n \times n$ matrix D such that $AD = I_n$.

l. A^T is an invertible matrix.

First, we need some notation. If the truth of statement (a) always implies that statement (j) is true, we say that (a) implies (j) and write $(a) \implies (j)$. The proof will establish the "circle" of implications shown in Fig. 12.1. If any one of these five statements is true, then so are the others. Finally, the proof will link the remaining statements of the theorem to the statements in this circle (see Figure 12.2).

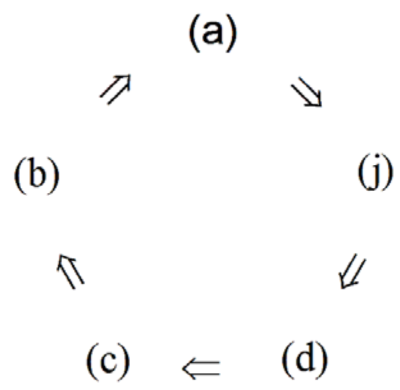


Fig. 12.1. A circle of implications

Proof. If statement (a) is true, then A^{-1} works for C in (j), so $(a) \Rightarrow (j)$.

Next, $(j) \Rightarrow (d)$ as if $\mathbf{x}$ satisfies $A\mathbf{x} = \mathbf{0}$, then $CAx = C\mathbf{0} = \mathbf{0}$ and so $I_n\mathbf{x} = 0$ and $\mathbf{x} = 0$. This shows that the equation $A\mathbf{x} = 0$ has no free variables.

Also, $(d) \Rightarrow (c)$.

Suppose for this that A is $n \times n$ and the equation $A\mathbf{x} = 0$ has only the trivial solution. Then there are no free variables in this equation, and so A has n pivot columns. Since A is square and the n pivot positions must be in different rows, the pivots in an echelon form of A must be on the main diagonal. Hence A is row equivalent to the $n \times n$ identity matrix. If A is square and has n pivot positions, then the pivots must lie on the main diagonal, in which case the reduced echelon form of A is I_n: Thus $(c) \Rightarrow (b)$.

Also, $(b) \Rightarrow (a)$ by Theorem 11.8 in Chapter 11. This completes the circle in Fig. 12.1.

Next, $(a) \Rightarrow (k)$ because A^{-1} works for D.

Also, $(k) \Rightarrow (g)$ as from $AD = I_m$ it follows that for any $\mathbf{b} \in \mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution (think about the equation $AD\mathbf{x} = \mathbf{b}$) .

Similarly $(g) \Rightarrow (a)$ as A is row equivalent to I. So (k) and (g) are linked to the circle.

Further, (g), (h), and (i) are equivalent for any matrix. Thus, (h) and (i) are linked through (g) to the circle.

Since (d) is linked to the circle, so are (e) and (f), because (d), (e), and (f) are all equivalent for any matrix A.

Finally, $(a) \Rightarrow (l)$ by Theorem 11.4 in Chapter 11, and $(l) \Rightarrow (a)$ by the same theorem

with A and A^T interchanged. ■

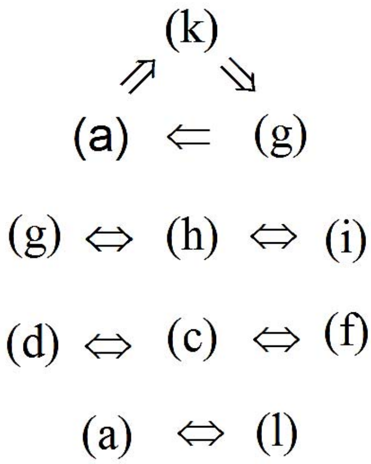


Fig. 12.2. Links from the remaining statements of the theorem to the statements in the circle presented in Figure 12.1

The power of the Invertible Matrix Theorem lies in the connections it provides among so many important concepts, such as linear independence of columns of a matrix A and the existence of solutions to equations of the form $A\mathbf{x} = \mathbf{b}$. It should be emphasized, however, that the Invertible Matrix Theorem applies *only to square matrices*. For example, if the columns of a 4×3 matrix are linearly independent, we cannot use the Invertible Matrix Theorem to conclude anything about the existence or nonexistence of solutions to equations of the form $A\mathbf{x} = \mathbf{b}$.

Example 12.2 *Use the Invertible Matrix Theorem to determine if A is invertible, where*

$$A = \begin{bmatrix} 1 & -3 & 0 \\ -4 & 11 & 1 \\ 2 & 7 & 3 \end{bmatrix}.$$

Solution:

$$A = \begin{bmatrix} 1 & -3 & 0 \\ -4 & 11 & 1 \\ 2 & 7 & 3 \end{bmatrix} \sim \cdots \sim \underbrace{\begin{bmatrix} 1 & -3 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 16 \end{bmatrix}}_{\text{3 pivots positions}}$$

Circle correct conclusion: Matrix A *is* / *is not* invertible. □

Example 12.3 *Suppose H is a 5×5 matrix and suppose there is a vector $\mathbf{v}$ in $\mathbb{R}^5$ which is not a linear combination of the columns of H. What can you say about the number of solutions to $H\mathbf{x} = \mathbf{0}$?*

Solution: Since $\mathbf{v}$ in $\mathbb{R}^5$ is not a linear combination of the columns of H, the columns of H do not ____________ $\mathbb{R}^5$.
So by the Invertible Matrix Theorem, $H\mathbf{x} = \mathbf{0}$ has

____.

□

12.2 Invertible Linear Transformations

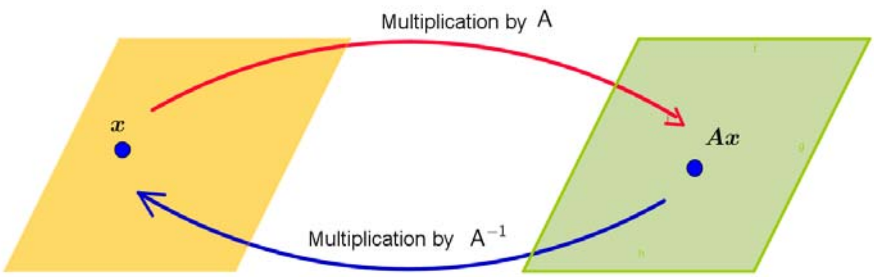
For an invertible matrix A,

$$A^{-1}A\mathbf{x} = \mathbf{x} \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

and

$$AA^{-1}\mathbf{x} = \mathbf{x} \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^n.$$

Picture:



Definition 12.4 *A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be **invertible** if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$S(T(\mathbf{x})) = \mathbf{x} \textit{ for all } \mathbf{x} \textit{ in } \mathbb{R}^n$$

and

$$T(S(\mathbf{x})) = \mathbf{x} \textit{ for all } \mathbf{x} \textit{ in } \mathbb{R}^n.$$

Theorem 12.5 *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let A be the standard matrix for T. Then T is invertible if and only if A is an invertible matrix. In that case, the linear transformation S given by $S(\mathbf{x}) = A^{-1}\mathbf{x}$ is the unique function satisfying*

$$S(T(\mathbf{x})) = \mathbf{x} \textit{ for all } \mathbf{x} \textit{ in } \mathbb{R}^n$$

and

$$T(S(\mathbf{x})) = \mathbf{x} \textit{ for all } \mathbf{x} \textit{ in } \mathbb{R}^n.$$

13

Introduction to Determinants

With each square matrix, it is possible to associate a real number called the determinant of the matrix. The value of this number will tell us whether the matrix is singular.

13.1 Preliminaries

Notation: A_{ij} is the matrix obtained from matrix A by deleting the i-th row and j-th column of A.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \qquad A_{23} = \begin{bmatrix} \quad & \quad & \quad \\ \\ \\ \\ \end{bmatrix}$$

□

Recall that $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$ and we let $\det [a] = a$.

Definition 13.1 *For $n \geq 2$, the **determinant** of an $n \times n$ matrix $A = [a_{ij}]$ is given by*

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n}$$

$$= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$$

Example 13.2 *Compute the determinant of*

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

Solution:

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

$= \underline{\hspace{10cm}} = \underline{\hspace{2cm}}$

□

Common notation: $\det \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix}$.

So

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

The $(\mathbf{i}, \mathbf{j})$**-cofactor** of A is the number C_{ij} where $C_{ij} = (-1)^{i+j} \det A_{ij}$.

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = 1C_{11} + 2C_{12} + 0C_{13}$$

(cofactor expansion across row 1)

Theorem 13.3 *The determinant of an $n \times n$ matrix A can be computed by a cofactor expansion across any row or down any column:*

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \quad \textit{(expansion across row } i\textit{)}$$

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \qquad \textit{(expansion down column } j\textit{)}$$

Use a matrix of signs to determine $(-1)^{i+j}$

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Exercise 13.1 *Compute the determinant of* $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{bmatrix}$ *using cofactor expansion down column* 3.

Solution:

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

□

Example 13.4 *Compute the determinant of*

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

Solution:

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

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

$$= 1 \cdot 2 \begin{vmatrix} 2 & 1 \\ 3 & 5 \end{vmatrix} = 14$$

Remark 13.5 *Method of cofactor expansion is not practical for large matrices.*

Triangular Matrices:

$$\begin{bmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ 0 & 0 & \ddots & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * \end{bmatrix} \qquad \begin{bmatrix} * & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & \ddots & 0 & 0 \\ * & * & \cdots & * & 0 \\ * & * & \cdots & * & * \end{bmatrix}$$

(upper triangular) (lower triangular)

Theorem 13.6 *If A is a triangular matrix, then* $\det A$ *is the product of the main diagonal entries of A.*

$$\begin{vmatrix} 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & -3 & 5 \\ 0 & 0 & 0 & 4 \end{vmatrix} = \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} = -24$$

□

13.2 Properties of Determinants

Theorem 13.7 *Let A be a square matrix.*

a) If a multiple of one row of A is added to another row of A to produce a matrix B, then $\det A = \det B$.

b) If two rows of A are interchanged to produce B, then $\det B = -\det A$.

c) If one row of A is multiplied by k to produce B, then $\det B = k \cdot \det A$.

Example 13.8 *Compute*

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 0 & 0 \\ 2 & 7 & 6 & 10 \\ 2 & 9 & 7 & 11 \end{vmatrix}.$$

Solution:

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

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

Theorem 13.7 (c) indicates that $\begin{vmatrix} * & * & * \\ -2k & 5k & 4k \\ * & * & * \end{vmatrix} = k \begin{vmatrix} * & * & * \\ -2 & 5 & 4 \\ * & * & * \end{vmatrix}$.

Example 13.9 *Compute*

$$\begin{vmatrix} 2 & 4 & 6 \\ 5 & 6 & 7 \\ 7 & 6 & 10 \end{vmatrix}.$$

Solution:

$$\begin{vmatrix} 2 & 4 & 6 \\ 5 & 6 & 7 \\ 7 & 6 & 10 \end{vmatrix} = 2 \begin{vmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 7 & 6 & 10 \end{vmatrix} = 2 \begin{vmatrix} 1 & 2 & 3 \\ 0 & -4 & -8 \\ 0 & -8 & -11 \end{vmatrix}$$

$$= 2(-4) \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -8 & -11 \end{vmatrix} = 2(-4) \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{vmatrix}$$

$$= 2(-4)(1)(1)(5) = -40$$

□

Example 13.10 *Compute*

$$\begin{vmatrix} 2 & 3 & 0 & 1 \\ 4 & 7 & 0 & 3 \\ 7 & 9 & -2 & 4 \\ 1 & 2 & 0 & 4 \end{vmatrix}$$

using a combination of row reduction and cofactor expansion.

Solution:

$$\begin{vmatrix} 2 & 3 & 0 & 1 \\ 4 & 7 & 0 & 3 \\ 7 & 9 & -2 & 4 \\ 1 & 2 & 0 & 4 \end{vmatrix} = -2\begin{vmatrix} 2 & 3 & 1 \\ 4 & 7 & 3 \\ 1 & 2 & 4 \end{vmatrix} = -2\begin{vmatrix} 2 & 3 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 4 \end{vmatrix}$$

$$= 2\begin{vmatrix} 2 & 3 & 1 \\ 1 & 2 & 4 \\ 0 & 1 & 1 \end{vmatrix} = -2\begin{vmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -2\begin{vmatrix} 1 & 2 & 4 \\ 0 & -1 & -7 \\ 0 & 1 & 1 \end{vmatrix}$$

$$= -2\begin{vmatrix} 1 & 2 & 4 \\ 0 & -1 & -7 \\ 0 & 0 & -6 \end{vmatrix} = -2\,(1)\,(-1)\,(-6) = -12.$$

□

Suppose A has been reduced to

$$U = \begin{bmatrix} \blacksquare & * & * & \cdots & * \\ 0 & \blacksquare & * & \cdots & * \\ 0 & 0 & \blacksquare & \cdots & * \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \blacksquare \end{bmatrix}$$

by row replacements and row interchanges, then

$$\det A = \begin{cases} (-1)^r \begin{pmatrix} \text{product of} \\ \text{pivots in } U \end{pmatrix} & \text{when } A \text{ is invertible} \\ 0 & \text{when } A \text{ is not invertible} \end{cases}$$

Theorem 13.11 *A square matrix is invertible if and only if* $\det A \neq 0$.

Theorem 13.12 *If A is an $n \times n$ matrix, then* $\det A^T = \det A$.

Partial proof
(2×2 case)

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \qquad \text{and}$$
$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix}^T = \det\begin{bmatrix} a & c \\ b & d \end{bmatrix} = ad - bc$$
$$\Rightarrow \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \det\begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

(3×3 case)

$$\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$
$$\det\begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix} = a\begin{vmatrix} e & h \\ f & i \end{vmatrix} - b\begin{vmatrix} d & g \\ f & i \end{vmatrix} + c\begin{vmatrix} d & g \\ e & h \end{vmatrix}$$
$$\Rightarrow \det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \det\begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}.$$

Implications of Theorem 13.12?
Theorem 13.7 still holds if the word *row* is replaced with ______________________.

Theorem 13.13 (Multiplicative Property) *For $n \times n$ matrices A and B, $\det(AB) = (\det A)(\det B)$.*

Example 13.14 *Compute $\det A^3$ if $\det A = 5$.*

Solution:

$$\det A^3 = \det(AAA) = (\det A)(\det A)(\det A).$$

$$= \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_\_.$$

□

Example 13.15 *For $n \times n$ matrices A and B, show that A is singular if $\det B \neq 0$ and $\det AB = 0$.*

Solution: *Since*

$$(\det A)(\det B) = \det AB = 0$$

and

$$\det B \neq 0,$$

then $\det A = 0$. Therefore A is singular. □

13.3 Cramer's rule

Cramer's rule is needed in a variety of theoretical calculations. For instance, it can be used to study how the solution of $A\mathbf{x} = \mathbf{b}$ is affected by changes in the entries of $\mathbf{b}$. However, the formula is inefficient for hand calculations, except for 2×2 or perhaps 3×3 matrices

For any $n \times n$ matrix A and any $\mathbf{b}$ in $\mathbb{R}^n$, let $A_i(\mathbf{b})$ be the matrix obtained from A by replacing column i by the vector $\mathbf{b}$.

$$A_i(\mathbf{b}) = \begin{bmatrix} \mathbf{a}_1 & \cdots & \underset{\text{col } i}{\underset{\uparrow}{\mathbf{b}}} & \cdots & \mathbf{a}_n \end{bmatrix}$$

Theorem 13.16 (Cramer's Rule) *Let A be an invertible $n \times n$ matrix. For any any $\mathbf{b}$ in $\mathbb{R}^n$, the unique solution $\mathbf{x}$ of $A\mathbf{x} = b$ has entries given by*

$$\mathbf{x}_i = \frac{\det A_i(\mathbf{b})}{\det A}, \qquad i = 1, 2, ..., n. \tag{13.1}$$

Proof. Denote the columns of A by $\mathbf{a}_1, ..., \mathbf{a}_n$ and the columns of the $n \times n$ identity matrix I by $\mathbf{e}_1, ..., \mathbf{e}_n$. If $A\mathbf{x} = \mathbf{b}$, the definition of matrix multiplication shows that

$$\begin{aligned} AI_i(\mathbf{x}) &= A\begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{x} & \cdots & \mathbf{a}_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{a}_1 & \cdots & A\mathbf{x} & \cdots & A\mathbf{a}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{b} & \cdots & \mathbf{a}_n \end{bmatrix} = A_i(\mathbf{b}). \end{aligned}$$

By the multiplicative property of determinants,

$$(\det A)(\det I_i(\mathbf{x})) = \det A_i(\mathbf{b})$$

(see Fig. 13.1). The second determinant on the left is simply $\mathbf{x}_i$. (Make a cofactor expansion along the i-th row.) Hence

$$(\det A)\,\mathbf{x}_i = \det A_i(\mathbf{b})$$

This proves (13.1) because A is invertible and $\det A \neq 0$. ■

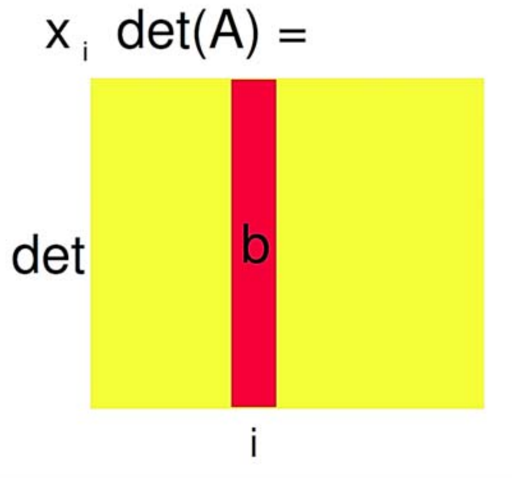


Fig. 13.1. Cramer's Rule explined.

The inverse of a matrix. Because the columns of A^{-1} are solutions of $A\mathbf{x} = \mathbf{e}_i$, where $\mathbf{e}_j$ are basis vectors, Cramer's rule together with the Laplace expansion gives the formula:

$$\left[A^{-1}\right]_{i,j} = (-1)^{i+j} \frac{\det(A_{ji})}{\det(A)} \qquad (13.2)$$

The matrix of cofactors $C_{ij} = (-1)^{i+j} \det(A_{ji})$ is called the **classical adjoint**[1] (or **adjugate matrix**) of A, denoted by adj A. Note the change $ij \to ji$. Don't confuse the classical adjoint with the transpose A^T which is sometimes also called the adjoint. Thus

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & \\ \vdots & & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

Warning: Note that the subscripts on C_{ji} are the reverse of (i, j).]

[1] The term *adjoint* also has another meaning in advanced texts on linear transformations.

14

Eigenvectors and Eigenvalues

14.1 An Introduction to Eigenproblems

The basic concepts presented here - *eigenvectors* and - are useful throughout pure and applied mathematics. Eigenvalues are also used to study difference equations and *continuous* dynamical systems. They provide critical information in engineering design, and they arise naturally in such fields as physics and chemistry.

Definition 14.1 *An* ***eigenvector*** *of an $n \times n$ matrix A is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. A scalar λ is called an* ***eigenvalue*** *of A if there is a nontrivial solution $\mathbf{x}$ of $A\mathbf{x} = \lambda\mathbf{x}$; such an $\mathbf{x}$ is called an eigenvector corresponding to λ.*

Example 14.2 *Let $A = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Examine the images of $\mathbf{u}$ and $\mathbf{v}$ under multiplication by A.*

Solution:

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

$$A\mathbf{v} = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \end{bmatrix} \neq \lambda\mathbf{v}$$

a) $\mathbf{u}$ is called an *eigenvector* of A.

b) $\mathbf{v}$ is not an eigenvector of A since $A\mathbf{v}$ is not a multiple of $\mathbf{v}$ (see Figure 14.1).

Example 14.3 *Show that 4 is an eigenvalue of $A = \begin{bmatrix} 0 & -2 \\ -4 & 2 \end{bmatrix}$ and find the corresponding eigenvectors.*

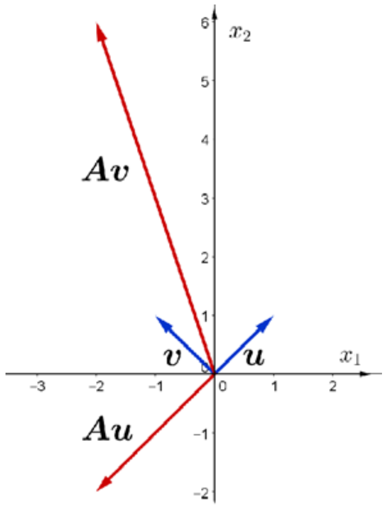


Fig. 14.1. $A\mathbf{u} = -2\mathbf{u}$, but $A\mathbf{v} \neq \lambda\mathbf{v}$

Solution: Scalar 4 is an eigenvalue of A if and only if $A\mathbf{x} = 4\mathbf{x}$ has a nontrivial solution.

$$A\mathbf{x}-4\mathbf{x} = \mathbf{0}$$

$$A\mathbf{x}-4\,(\_\_\_\_)\,\mathbf{x} = \mathbf{0}$$

$$(A-4I)\,\mathbf{x} = \mathbf{0}.$$

To solve $(A-4I)\,\mathbf{x} = \mathbf{0}$, we need to find $A-4I$ first:

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

Now solve $(A-4I)\,\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} -4 & -2 & 0 \\ -4 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} \frac{-1}{2}x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} \frac{-1}{2} \\ 1 \end{bmatrix}.$$

Each vector of the form $x_2 \begin{bmatrix} \frac{-1}{2} \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda = 4$ (see Figure 14.2.) $\square$

Warning: The method just used to find eigen*vectors cannot* be used to find eigen*values*.

Definition 14.4 *The set of all solutions to* $(A-\lambda I)\,\mathbf{x} = \mathbf{0}$ *is called the* ***eigenspace*** *of* A *corresponding to* λ.

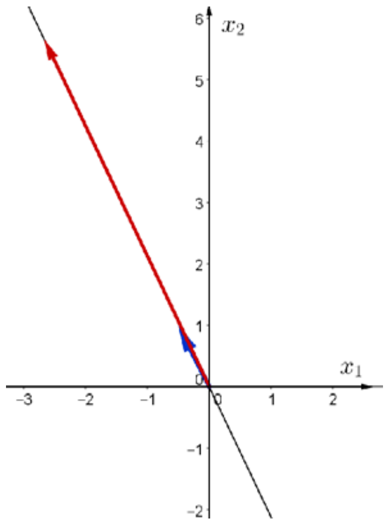


Fig. 14.2. Eigenspace for $\lambda = 4$

Example 14.5 *Let*

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

An eigenvalue of A is $\lambda = 2$. Find a basis for the corresponding eigenspace.

Solution:

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

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

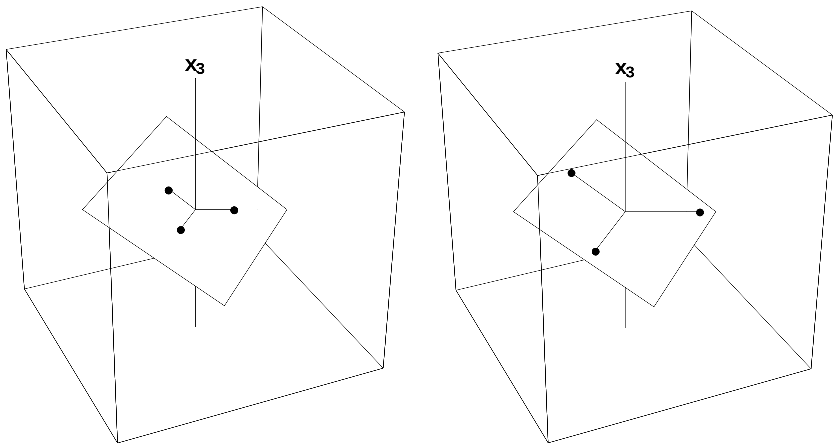
Augmented matrix for $(A - 2I)\,\mathbf{x} = \mathbf{0}$:

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

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 + x_3 \\ x_2 \\ x_3 \end{bmatrix} = \_\_\_\_ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \_\_\_\_ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

So a basis for the eigenspace corresponding to $\lambda = 2$ is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$



Effects of Multiplying Vectors in Eigenspaces for $\lambda = 2$ by A

Example 14.6 *Suppose λ is eigenvalue of A. Determine an eigenvalue of A^2 and A^3. In general, what is an eigenvalue of A^n?*

Solution: Since λ is eigenvalue of A, there is a nonzero vector $\mathbf{x}$ such that

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

Then

$$\_\_\_A\mathbf{x} = \_\_\_\lambda\mathbf{x}$$

$$A^2\mathbf{x} = \lambda A\mathbf{x}$$
$$A^2\mathbf{x} = \lambda\_\_\_\mathbf{x}$$
$$A^2\mathbf{x} = \lambda^2\mathbf{x}$$

Therefore λ^2 is an eigenvalue of A^2.
Show that λ^3 is an eigenvalue of A^3:

$$\_\_\_A^2\mathbf{x} = \_\_\_\lambda^2\mathbf{x}$$
$$A^3\mathbf{x} = \lambda^2 A\mathbf{x}$$
$$A^3\mathbf{x} = \lambda^3\mathbf{x}$$

Therefore λ^3 is an eigenvalue of A^3.

In general, ______ is an eigenvalue of A^n.

Theorem 14.7 *The eigenvalues of a triangular matrix are the entries on its main diagonal.*

Proof for the 3×3 Upper Triangular Case: Let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}.$$

and then

$$A - \lambda I = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$
$$= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix}.$$

By definition, λ is an eigenvalue of A if and only if $(A - \lambda I)\,\mathbf{x} = \mathbf{0}$ has a nontrivial solution. This occurs if and only if $(A - \lambda I)\,\mathbf{x} = \mathbf{0}$ has a free variable.

When does this occur?

Theorem 14.8 *If* $\mathbf{v}_1, \ldots, \mathbf{v}_r$ *are eigenvectors that correspond to distinct eigenvalues* $\lambda_1, \ldots, \lambda_r$ *of an* $n \times n$ *matrix* A*, then* $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ *is a linearly independent set.*

Proof. Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly dependent. Since $\mathbf{v}_1$ is nonzero, one of the vectors in the set is a linear combination of the preceding vectors. Let p be the least index such that $\mathbf{v}_{p+1}$ is a linear combination of the preceding (linearly independent) vectors. Then there exist scalars $c_1, ..., c_p$ such that

$$c_1\mathbf{v}_1 + ..., c_p\mathbf{v}_p = \mathbf{v}_{p+1} \tag{14.1}$$

Multiplying both sides of (14.1) by A and using the fact that $Av_k = \lambda_k\mathbf{v}_k$ for each k, we obtain

$$c_1A\mathbf{v}_1 + ... + c_pA\mathbf{v}_p = A\mathbf{v}_{p+1}$$

$$c_1\lambda_1\mathbf{v}_1 + ..., c_p\lambda_p\mathbf{v}_p = \lambda_{p+1}\mathbf{v}_{p+1}. \tag{14.2}$$

Multiplying both sides of (14.1) by λ_{p+1} and subtracting the result from (14.2), we have

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + ..., c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = \mathbf{0} \tag{14.3}$$

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is linearly independent, the weights in (14.3) are all zero. But none ofthe factors $(\lambda_i - \lambda_{p+1})$ are zero, because the eigenvalues are distinct. Hence $c_i = 0$ for $i = 1, ..., p$. But then (14.1) says that $\mathbf{v}_{p+1} = 0$, which is impossible. Hence $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ cannot be linearly dependent and therefore must be linearly independent. ■

15

The Characteristic Equation

15.1 Review

$$A \quad \mathbf{x} \quad = \quad \lambda \quad \mathbf{x}$$

Find eigenvectors $\mathbf{x}$ by solving $(A - \lambda I)\,\mathbf{x} = \mathbf{0}$.
How do we find the eigenvalues λ**?**

$$\mathbf{x} \text{ must be nonzero}$$
$$\Downarrow$$
$$(A - \lambda I)\,\mathbf{x} = \mathbf{0} \text{ must have nontrivial solutions}$$
$$\Downarrow$$
$$(A - \lambda I) \text{ is not invertible}$$
$$\Downarrow$$
$$\det(A - \lambda I) = 0$$

(called the *characteristic equation*)

$$\boxed{\text{Solve } \det(A - \lambda I) = 0 \text{ for } \lambda \text{ to find the eigenvalues.}}$$

Characteristic polynomial: $\det(A - \lambda I)$

Characteristic equation: $\det(A - \lambda I) = 0$

Example 15.1 *Find the eigenvalues of*

$$A = \begin{bmatrix} 0 & 1 \\ -6 & 5 \end{bmatrix}.$$

Solution: Since

$$A - \boldsymbol{\lambda} I = \begin{bmatrix} 0 & 1 \\ -6 & 5 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -\lambda & 1 \\ -6 & 5 - \lambda \end{bmatrix},$$

the equation $\det(A - \boldsymbol{\lambda} I) = 0$ becomes

$$-\lambda(5-\lambda)+6=0$$

$$\lambda^2-5\lambda+6=0$$

Factor:

$$(\lambda-2)(\lambda-3)=0.$$

So the eigenvalues are 2 and 3. □

For a 3×3 matrix or larger, recall that a determinant can be computed by cofactor expansion.

Example 15.2 *Find the eigenvalues of*

$$A=\begin{bmatrix}1 & 2 & 1\\ 0 & -5 & 0\\ 1 & 8 & 1\end{bmatrix}.$$

Solution:

$$A-\boldsymbol{\lambda}I=\begin{bmatrix}1-\_\_\_\_ & 2 & 1\\ 0 & -5-\_\_\_\_ & 0\\ 1 & 8 & 1-\_\_\_\_\end{bmatrix}$$

$$\det(A-\boldsymbol{\lambda}I)=\begin{vmatrix}1-\lambda & 2 & 1\\ 0 & -5-\lambda & 0\\ 1 & 8 & 1-\lambda\end{vmatrix}=(-5-\lambda)\begin{vmatrix}1-\lambda & 1\\ 1 & 1-\lambda\end{vmatrix}$$

$$=(-5-\lambda)\left[(1-\lambda)^2-1\right]=(-5-\lambda)\left[1-2\lambda+\lambda^2-1\right]$$

$$=(-5-\lambda)\left[-2\lambda+\lambda^2\right]=-(5+\lambda)\lambda[-2+\lambda]=0$$

$$\Rightarrow \lambda=-5,\ 0,\ 2$$

□

Theorem 15.3 *(The Invertible Matrix Theorem - continued) Let A be an $n\times n$ matrix. Then A is invertible if and only if:*

s) The number 0 is not an eigenvalue of A.

t) $\det A \neq 0$

Recall that if B is obtained from A by a sequence of row replacements or interchanges, but without scaling, then $\det A = (-1)^r \det B$, where r is the number of row interchanges.

Suppose the echelon form U is obtained from A by a sequence of row replacements or interchanges, but without scaling.

$$A \sim U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & u_{nn} \end{bmatrix}$$

The **determinant** of A, written $\det A$, is defined as follows:

$$\det A = \begin{cases} (-1)^r \cdot \left(\begin{array}{l} \text{product of} \\ \text{pivots in } U \end{array} \right), & \text{when } A \text{ is invertible} \\ 0, & \text{when } A \text{ is not invertible} \end{cases}$$

(r is the number of row interchanges)

Example 15.4 *Find the eigenvalues of*

$$A = \begin{bmatrix} 3 & 2 & 3 \\ 0 & 6 & 10 \\ 0 & 0 & 2 \end{bmatrix}.$$

Solution:

$$\det(A - \lambda I) = \det \begin{bmatrix} 3-\lambda & 2 & 3 \\ 0 & 6-\lambda & 10 \\ 0 & 0 & 2-\lambda \end{bmatrix}$$

Characteristic equation:

$$(\qquad\qquad)(\qquad\qquad)(\qquad\qquad) = 0.$$

eigenvalues: _____, _____, _____

□

Definition 15.5 *The **(algebraic) multiplicity** of an eigenvalue is its multiplicity as a root of the characteristic equation.*

Example 15.6 *Find the characteristic polynomial of*

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 5 & 3 & 0 & 0 \\ 9 & 1 & 3 & 0 \\ 1 & 2 & 5 & -1 \end{bmatrix}$$

and then find all the eigenvalues and the algebraic multiplicity of each eigenvalue.

Solution:

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 0 & 0 & 0 \\ 5 & 3-\lambda & 0 & 0 \\ 9 & 1 & 3-\lambda & 0 \\ 1 & 2 & 5 & -1-\lambda \end{vmatrix}$$

$$= (2-\lambda)(3-\lambda)(3-\lambda)(-1-\lambda) = 0$$

eigenvalues: _____, _____, _____

□

15.2 Similarity

Numerical methods for finding approximating eigenvalues are based upon Theorem 15.7 to be described shortly. **??**

For $n \times n$ matrices A and B, we say the A is **similar** to B if there is an invertible matrix P such that

$$P^{-1}AP = B \qquad \text{or equivalently,} \qquad A = PBP^{-1}.$$

Theorem 15.7 *If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).*

Proof. If $B = P^{-1}AP$, then

$$\det(B - \lambda I) = \det\left[P^{-1}AP - P^{-1}\lambda IP\right] = \det\left[P^{-1}(A - \lambda I)P\right]$$
$$= \det P^{-1} \cdot \det(A - \lambda I) \cdot \det P = \det(A - \lambda I).$$

■

15.3 Application to Markov Chains

Example 15.8 *Consider the migration matrix* $M = \begin{bmatrix} .95 & .90 \\ .05 & .10 \end{bmatrix}$ *and define* $\mathbf{x}_{k+1} = M\mathbf{x}_k$. *It can be shown that*

$$\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$$

converges to a steady state vector $\mathbf{x} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$. *Why?*

The answer lies in examining the corresponding eigenvectors.
Solution: First we find the eigenvalues:

$$\det(M - \lambda I) = \det\left(\begin{bmatrix} .95 - \lambda & .90 \\ .05 & .10 - \lambda \end{bmatrix}\right) = \lambda^2 - 1.05\lambda + 0.05$$

So solve

$$\lambda^2 - 1.05\lambda + 0.05 = 0$$

By factoring

$$\lambda = 0.05,\ \lambda = 1$$

It can be shown that the eigenspace corresponding to $\lambda = 1$ is span$\{\mathbf{v}_1\}$ where $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and the eigenspace corresponding to $\lambda = 0.05$ is span$\{\mathbf{v}_2\}$ where $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Note that

$$M\mathbf{v}_1 = \mathbf{v}_1,$$

and so $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$ is our steady state vector.

Then for a given vector $\mathbf{x}_0$,

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$$
$$\mathbf{x}_1 = M\mathbf{x}_0 = M(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1M\mathbf{v}_1 + c_2M\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2(0.05)\mathbf{v}_2$$

$$\mathbf{x}_2 = M\mathbf{x}_1 = M\left(c_1\mathbf{v}_1 + c_2\left(0.05\right)\mathbf{v}_2\right) = c_1 M\mathbf{v}_1 + c_2\left(0.05\right)M\mathbf{v}_2 =$$
$$c_1\mathbf{v}_1 + c_2\left(0.05\right)^2\mathbf{v}_2$$
$$\text{and in general}$$
$$\mathbf{x}_k = c_1\mathbf{v}_1 + c_2\left(0.05\right)^k\mathbf{v}_2$$
$$\text{and so } \lim_{k\to\infty}\mathbf{x}_k = \lim_{k\to\infty}\left(c_1\mathbf{v}_1 + c_2\left(0.05\right)^k\mathbf{v}_2\right) = c_1\mathbf{v}_1$$

and this is the steady state when $c_1 = \frac{1}{2}$.

16

Appendixes

16.1 Greek letters used in mathematics, science, and engineering

The Greek letter forms used in mathematics are often different from those used in Greek-language text: they are designed to be used in isolation, not connected to other letters, and some use variant forms which are not normally used in current Greek typography. The table below shows Greek letters rendered in TEX

Table 16.1. Greek letters used in mathematics

α		alpha	ν		nu
β		beta	ξ	Ξ	xi
γ	Γ	gamma	π	Π	pi
δ	Δ	delta	ρ		rho
ϵ		epsilon	σ	Σ	sigma
ζ		zeta	τ		tau
η		eta	υ		upsilon
θ	Θ	theta	ϕ	Φ	phi
ι		iota	χ		chi
κ		kappa	ψ	Ψ	psi
λ	Λ	lambda	ω	Ω	omega
μ		mu	$\dagger$		dagger

TEXis a typesetting system designed and mostly written by Donald Knuth at Stanford and released in 1978.

Together with the Metafont language for font description and the Computer Modern family of typefaces, TeX was designed with two main goals in mind: to allow anybody to produce high-quality books using a reasonably minimal amount of effort, and to provide a system that would give exactly the same results on all computers, now and in the future.

Bibliography

[1] S. Andrilli and D. Hecker (2010). *Elementary Linear Algebra*, Elsevier, Amsterdam.

[2] H. Anton and Ch. Rorres (2005). *Elementary Linear Algebra. Applications version.* John Wiley & Sons, New York.

[3] M. Anthony and M. Harvey (2012). *Linear Algebra: Concepts and Methods,* Cambridge University Press, Cambridge.

[4] S. Axler (1997). *Linear Algebra Done Right,* Springer, New York.

[5] R.A.Beezer (2013) *A First Course in Linear Algebra*, Congruent Press, Washington.

[6] G. Birkhoff and S. MacLane (1997). *A Survey of Modern Algebra (4th ed.),* Macmillan, New York.

[7] O. Bretscher (2008). Linear Algebra with Applications (4th. ed.), Pearson, Boston.

[8] J.W. Brown and R.V. Churchill (2013). *Complex Variables and Applications*, McGraw-Hill, New York.

[9] D. Cherney, T. Denton and A. Waldron (2013). *Linear Algebra,* https://www.math.ucdavis.edu/~linear/

[10] G. Farin and D. Hansford (2005). *Practical Linear Algebra. A Geometry Toolbox*, A.K. Peters, Wellesley, Massachusetts.

[11] R. Hammack (2013). *Book of Proof*, Richard Hammack (publisher), Richmond, Virginia.

[12] J. Hefferton, *Linear Algebra*, http://joshua.smcvt.edu/linearalgebra

[13] J.M. Howe (2003), *Complex Analysis,* Springer, London.

[14] D.C. Lay (2012), *Linear Algebra and Its Applications,* Addison-Wesley, Boston.

[15] D. McMahon (2006). *Linear Algebra Demystified,* McGraw-Hill, New York.

[16] D. McMahon (2006). *Complex Variables Demystified,* McGraw-Hill, New York.

[17] C.D. Meyer (2000). *Matrix Analysis and Applied Linear Algebra,* SIAM, Philadelphia.

[18] T. S. Shores (2000). *Applied Linear Algebra and Matrix Analysis,* Springer, Berlin.

[19] G. Strang (2009). *Introduction to Linear Algebra (4th ed.),* Wellesley-Cambridge Press, Wellesley MA.

[20] G. Strang (2006). *Linear Algebra and Its Applications (4th ed.),* Thomson Learning, Belmont Ca.

Index

Index