1 Vector and Matrix Algebra

Mathematics is the music of reason.

The world of ideas which [mathematics] discloses or illuminates, the contemplation of divine beauty and order which it induces, the harmonious connexion of its parts, the infinite hierarchy and absolute evidence of the truths with which it is concerned, these, and such like, are the surest grounds of the title of mathematics to human regard, and would remain unimpeached and unimpaired were the plan of the universe unrolled like a map at our feet, and the mind of man qualified to take in the whole scheme of creation at a glance. (James Joseph Sylvester)

Approximately 50 years before publication of this book, Neil Armstrong was the first human to step foot on the Moon and – thankfully – return safely to Earth. While the Apollo 11 mission that facilitated this amazing accomplishment in July of 1969 was largely symbolic from a scientific perspective, it represented the culmination of a decade of focused effort to create, develop, test, and implement a whole host of new technologies that were necessary to make such a feat possible. Many of these technologies are now fully integrated into modern life.

The march to the Moon during the 1960s was unprecedented in modern technological history for its boldness, aggressive time frame, scale, scientific and technological developments, and ultimate success. The state of the art in space travel in the early 1960s was to put a human in orbit 280 kilometers (km) (175 miles) above the Earth for several hours. It was in this context on September 12, 1962, that President Kennedy boldly – and some argued foolishly – set the United States on a course "to go to the Moon in this decade and do the other things, not because they are easy, but because they are hard." The Moon is 385,000 km (240,000 miles) from Earth – orbital flights were a mere baby step toward this aggressive goal. It was not that the technology was available and we simply needed to marshal the financial resources and political will to accomplish this bold task, it was not clear *if* the technology could be developed at all, let alone in the aggressive time frame proposed.

This problem appeared straightforward on paper. Newton's laws of motion had been articulated nearly 300 years before, and they had served remarkably well as the basis for an uncountable array of terrestrial applications and cosmological predictions. But to successfully land an object the size of a small truck on a moving body 385,000 km from its origin presented a whole host of issues. The mathematical model

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

in the form of Newton's laws could not be fully trusted to accurately navigate such extreme distances. Nor could noisy and inaccurate measurements made on board the spacecraft as well as back on Earth be relied upon either. How could a variety of measurements, each with various levels of reliability and accuracy, be combined with mathematical predictions of the spacecraft's trajectory to provide the best estimate of its location and determine the necessary course corrections to keep it on its precise trajectory toward its target – and back?

The obstacle was not that the equations were not correct, or that they could not be accurately calculated; the problem was that there were numerous opportunities for the spacecraft to deviate slightly from its intended path. Each of these could have led to large deviations from the target trajectory over such large distances. A mechanism was required to correct the trajectory of the spacecraft along its route, but that was the easy part. How does one determine precisely where you are when so far from, well, anything? How does one adjust for the inevitable errors in making such a *state estimate*?

While so-called *filtering* techniques were available at the time, they were either not accurate enough for such a mission or too computationally intensive for on-board computers. It turned out that the methodology to address just such a predicament had been published by Kalman (1960) only 18 months before Kennedy's famous speech. In fact, a small group of researchers from NASA Ames Research Center had been in contact with Kalman in the late 1950s to discuss his new method and how it could be used for midcourse navigation correction of the Apollo spacecraft.[1] Kalman's original method applied to linear estimation problems, whereas NASA scientists were faced with a nonlinear problem. The modifications made to the original approach came to be known as the "extended Kalman filter," which remains one of the dominant approaches for treating nonlinear estimation problems today.

While on-board Kalman filtering ended up only being used as a backup to ground-based measurements for midcourse navigation between the Earth and Moon, it was found to be essential for the rendezvous problem. The three Apollo astronauts traveled together in the Command Module until the spacecraft entered lunar orbit. The Lunar Module then detached from the Command Module with Neil Armstrong and Buzz Aldrin aboard, while Michael Collins stayed in the Command Module as it remained in lunar orbit. After their historic landing on the Moon's surface, the Lunar Module then had to take off from the lunar surface and rendezvous with the Command Module in lunar orbit.[2] This was the most difficult navigational challenge of the entire mission. Because both spacecraft were moving, it presented a unique navigational challenge to facilitate their rendezvous with the necessary precision to dock the two spacecraft and reunite the three history-making astronauts. Kalman filtering was used to ingest the model predictions and on-board measurements to obtain relative state estimates

[1] NASA Technical Memorandum 86847, "Discovery of the Kalman Filter as a Practical Tool for Aerospace and Industry," by L. A. McGee and S. F. Schmidt (1985).

[2] The picture on the cover of the book was taken by Michael Collins in the Command Module just before the rendezvous was executed.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

between the two orbiting spacecraft. The calculations were performed redundantly on both the Lunar and Command Modules' guidance computers for comparison.[3]

Before we can introduce state estimation in Section 11.9, there is a great deal of mathematical machinery that we need to explore. This includes material from this book on matrix methods as well as variational methods (see, for example, Cassel 2013). There are numerous additional examples of new technologies and techniques developed for the specific purpose of landing a man on the Moon – many of them depending upon the matrix, numerical, and optimization methods that are the subjects of this text.

1.1 Introduction

For most scientists and engineers, our first exposure to vectors and matrices is in the context of mechanics. Vectors are used to represent quantities, such as velocity or force, that have both a magnitude and direction in contrast to scalar quantities, such as pressure and temperature, that only have a magnitude. Likewise, matrices generally first appear when stress and strain tensors are introduced, once again in a mechanics setting. At first, such threedimensional vectors and matrices appear to be simply a convenient way to tabulate such quantities in an orderly fashion. In large part, this is true. In contrast to many areas of mathematics, for which certain operations would not be possible without it, vectors and matrices are not mathematical elements of necessity; rather they are constructs of convenience. It could be argued that there is not a single application in this book that *requires* matrix methods. However, this ubiquitous framework not only supplies a convenient way of representing large and complex data sets, it also provides a common formalism for their analysis; the topics contained in this book would be far more complex and confusing without the machinery of matrix methods.

The benefit of first being exposed to vectors and matrices through mechanics is that we naturally develop a strong geometric interpretation of them from the start. Because it appeals to our visual sensibilities, therefore, we hardly realize that we are learning the basics of *linear algebra*. In such a mechanics context, however, there is no reason to consider vectors larger than three dimensions corresponding to the three directions in our various coordinate systems that represent physical space. Unfortunately, our visual interpretation of vectors and matrices does not carry over to higher dimensions; we cannot even sketch a vector larger than three dimensions let alone impose a meaningful geometric interpretation. This is when linear algebra seems to lose its moorings in physically understandable reality and is simply a fun playground for mathematicians.

[3] MIT Report E-2411, "Apollo Navigation, Guidance, and Control: A Progress Report," by D. G. Hoag (1969). MIT Report R-649, "The Apollo Rendezvous Navigation Filter Theory, Description and Performance" (Volume 1), by E. S. Muller Jr. and P. M. Kachmar (1970).

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Given its roots in algebra, it is not surprising that mathematicians have long viewed linear algebra as an essential weapon in their arsenal. For those of us who have benefited from a course in linear algebra taught by a mathematician, complementing the geometrically rich and physically practical exposure in mechanics with the formal framework of operations and methods provides a strong, and frankly necessary, foundation for research and practice in almost all areas of science and engineering. However, the mathematician's discussion of "singular matrices," "null spaces," and "vector bases" often leaves us with the notion that linear algebra, beyond our initial mechanics-driven exposure, is of very little relevance to the scientist or engineer. On the contrary, the formalism of linear algebra provides the mathematical foundation for three of the most far-reaching and widely applicable "applications" of matrix methods in science and engineering, namely dynamical systems theory, numerical methods, and optimization. Together they provide the tools necessary to solve, analyze, and optimize large-scale, complex systems of practical interest in both research and industrial practice. As we will see, even the analysis and prediction of continuous systems governed by differential equations ultimately reduces to solution of a matrix problem. This is because more often than not, numerical methods must be used, which convert the continuous governing equations into a discrete system of algebraic equations.

Although "linear" algebra is strictly speaking a special case of algebra,[4] it is in many ways a dramatic extension of the algebra that we learn in our formative years. This is certainly the case with regard to the applications that matrix methods address for scientists and engineers. Because linear methods are so well developed mathematically, with their wide-ranging set of tools, it is often tempting to reframe nonlinear problems in such a way as to allow for the utilization of linear methods. In many cases, this can be formally justified; however, one needs to be careful to do so in a manner that is faithful to the true nature of the underlying system and the information being sought. This theme will be revisited throughout the text as many of our applications exhibit nonlinear behavior.

Whereas vectors and matrices arise in a wide variety of applications and settings, the mathematics of these constructs is the same regardless of where the vectors or matrices have their origin. We will focus in Part I on the mathematics, but with little emphasis on formalism and proofs, and mention or illustrate many of the applications in science and engineering. First, however, let us motivate the need for such mathematical machinery using two simple geometric scenarios. After introducing some basic definitions and algebraic operations, we will then return in Section 1.4.1 to introduce some additional applications of matrix problems common in science and engineering.

1.1.1 Equation of a Line

We know intuitively that the shortest distance between two points is a straight line. Mathematically, this is reflected in the fact that there is a single unique straight line

[4] Although algebra can be traced back to the ancient Babylonians and Egyptians, matrix algebra was not formalized until the middle of the nineteenth century by Arthur Cayley.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

that connects any two points. Consider the two points (x_1, y_1) and (x_2, y_2) in two dimensions. In order to determine the line,

$$a_0 + a_1 x = y,$$

passing through these two points, we could substitute the two points into this equation for the line as follows:

$$\begin{aligned} a_0 + a_1 x_1 &= y_1, \\ a_0 + a_1 x_2 &= y_2. \end{aligned} \tag{1.1}$$

Because the values of x and y are known for the two points, this is two equations for the two unknown constants a_0 and a_1, and we expect a unique solution. As will be shown in the next section, these coupled algebraic equations can conveniently be written in vector-matrix form as

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

In order to see how this corresponds to the system of equations (1.1), matrix multiplication will need to be defined. Working from left to right, we have a matrix that includes the x values multiplied by a vector containing the coefficients in the equation of the line set equal to a vector of the y values. In the present case, the matrix and right-hand-side vector are known, and a solution is sought for the coefficients a_0 and a_1 in the equation for the line. Performing the same exercise in three dimensions would result in the need to solve three equations for three unknowns, which again could be expressed in matrix form.

What if we have more than two points? Say that instead we have $N > 2$ such points. Certainly, it would not be possible to determine a single straight line that connects all N of the points (unless they so happen to all be collinear). Expressed as before, we would have N equations for the two unknown coefficients; this is called an *overdetermined system* as there are more equations than unknowns and no unique solution exists. While there is not a single line that connects all of the points in this general case, we could imagine that there is a single line that best represents the points as illustrated in Figure 1.1. This is known as *linear least-squares regression* and illustrates that there will be times when a "solution" of the system of equations is sought even when a unique solution does not exist. Least-squares methods will be taken up in Section 10.2 after we have covered the necessary background material.

1.1.2 Linear Transformation

As a second example, consider geometric transformations. When scientists and engineers communicate their ideas, theories, and results, they must always draw attention to their *reference frame*. For example, when describing the motion of a passenger walking down the aisle of an airplane in flight, is the description from the point of view of the passenger, the airplane, a fixed point on the Earth's surface, the center of the Earth, the center of the Sun, or some other point in the universe? Obviously, the

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

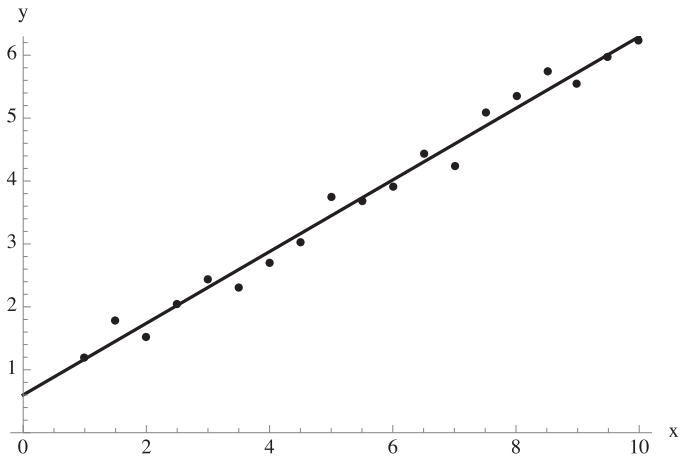


Figure 1.1 Best-fit line for $N = 19$ data points.

description will be quite different depending on which of these reference frames is used. Mathematically, we formally define the reference frame and how we will locate positions relative to it using a *coordinate system* or *vector basis*.[5] This is comprised of specifying the location of the coordinate system's origin – the location from which each coordinate and vector is measured – and each of the coordinate directions emanating from this origin.

Of course, one may specify a problem with respect to a different coordinate system than someone else, or different aspects of the problem may be best analyzed using different coordinate systems. In order to communicate information between the two coordinate systems, a *transformation* is necessary. For example, let us say that a point in the two-dimensional coordinate system X is given by $\mathbf{x}_1 = (2,5)$, and the same point with respect to another coordinate system Y is defined by $\mathbf{y}_1 = (-3,4)$. If this is a *linear* transformation, then we must be able to multiply $\mathbf{x}_1$ times something, say $\mathbf{A}$, to produce $\mathbf{y}_1$ in the form

$$\mathbf{A}\mathbf{x}_1 = \mathbf{y}_1. \tag{1.2}$$

Because both $\mathbf{x}_1$ and $\mathbf{y}_1$ involve two values, or coordinates, $\mathbf{A}$ is clearly not simply a single scalar value under general circumstances. Instead, we will need to include some combination of both coordinates to get from one coordinate system to the other. This could involve, for example, each coordinate of $\mathbf{y}_1$ being some linear combination of the two coordinates of $\mathbf{x}_1$, such that

[5] We generally call it a *coordinate system* when referring to a system with one, two, or three spatial dimensions, for which we can illustrate the coordinate system geometrically. For systems with more than three dimensions, that is, where the dimensions do not correspond to spatial coordinates, we use the more general *vector basis*, or simply *basis*, terminology. Consequently, a coordinate system is simply a vector basis for a two- or three-dimensional vector space.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

$$\begin{aligned} 2a + 5b &= -3, \\ 2c + 5d &= 4. \end{aligned} \tag{1.3}$$

If we suitably define multiplication, (1.3) can be written in the compact form of (1.2) if $\mathbf{A}$ is defined as

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

which is a *transformation matrix*.[6] It just remains to determine the constants a, b, c, and d that comprise the transformation matrix. What we have in (1.3) is a system of two coupled algebraic equations for these four unknown constants. Clearly, there is no unique solution as there are many, in fact infinite, ways to transform the single point $\mathbf{x}_1$ into the point $\mathbf{y}_1$. For example, we could simply translate the coordinate system or rotate it clockwise or counterclockwise to transform one point to the other.

A unique transformation is obtained if we supply an additional pair of *image* points in the two coordinate systems, say $\mathbf{x}_2 = (-1, 6)$ and $\mathbf{y}_2 = (8, -7)$. This results in the two additional equations

$$\begin{aligned} -a + 6b &= 8, \\ -c + 6d &= -7, \end{aligned} \tag{1.4}$$

which in compact form is

$$\mathbf{A}\mathbf{x}_2 = \mathbf{y}_2. \tag{1.5}$$

Note that the transformation matrix is the same in (1.2) and (1.5). Equations (1.3) and (1.4) now provide four equations for the four unknowns, and the transformation is unique.

Is there a way to represent the *system of linear algebraic equations* given by (1.3) and (1.4) in a convenient mathematical form? How do we solve such a system? Can we be sure that this solution is unique? Can we represent *linear transformations* in a general fashion for systems having any number of coordinates? These questions are the subject of this chapter. As we will see, the points and the coordinate directions with respect to which they are defined will be conveniently represented as *vectors*. The transformation will be denoted by a *matrix* as will be the coefficients in the system of linear algebraic equations to be solved for the transformation matrix. The unknowns in the transformation matrix will be combined to form the *solution vector*.

Such changes of coordinate system (basis) and the transformation matrices that accomplish them will be a consistent theme throughout the text. We will often have occasion to transform a problem into a more desirable basis in order to facilitate interpretation, expose features, diagnose attributes, or ease solution. We will return

[6] Given their ubiquity across so many areas of mathematics, science, and engineering, one might be surprised to learn that the term "matrix" was not coined until 1850 by James Joseph Sylvester. This is more than a century after such classical fields as complex variables, differential calculus, and variational calculus had their genesis.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

to this important topic in Section 1.8. Along the way, we will encounter numerous physical applications that lend themselves to similar mathematical representation as this geometric example despite their vastly different physical interpretations. As is so often the case in mathematics, this remarkable ability to unify many disparate applications within the same mathematical constructs and operations is what renders mathematics so essential to the scientist or engineer.

1.2 Definitions

Let us begin by defining vectors and matrices and several common types of matrices. One can think of a matrix as the mathematical analog of a table in a text or a spreadsheet on a computer.

Matrix: A *matrix* is an ordered arrangement of numbers, variables, or functions comprised of a rectangular grouping of elements arranged in rows and columns as follows:

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{bmatrix} = [A_{mn}] .$$

The size of the matrix is denoted by the number of rows, M, and the number of columns, N. We say that the matrix $\mathbf{A}$ is $M \times N$, which we read as "M by N." If $M = N$, then the matrix is said to be *square*. Each element A_{mn} in the matrix is uniquely identified by two subscripts, with the first, m, being its row and the second, n, being its column. Thus, $1 \leq m \leq M$ and $1 \leq n \leq N$. The elements A_{mn} may be real or complex numbers, variables, or functions.

The *main diagonal* of the matrix $\mathbf{A}$ is given by $A_{11}, A_{22}, \ldots, A_{MM}$ or A_{NN}; if the matrix is square, then $A_{MM} = A_{NN}$. Two matrices are said to be equal, that is $\mathbf{A} = \mathbf{B}$, if their sizes are the same and $A_{mn} = B_{mn}$ for all m and n.

Vector: A (column) *vector* is an $N \times 1$ matrix. For example,

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

The vector is said to be N-dimensional and can be considered a point in an N-dimensional coordinate system. By common convention, matrices are denoted by bold capital letters and vectors by bold lowercase letters.

Matrix Transpose (Adjoint): The *transpose* of matrix $\mathbf{A}$ is obtained by interchanging its rows and columns as follows:

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

$$\mathbf{A}^T = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{M1} \\ A_{12} & A_{22} & \cdots & A_{M2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1N} & A_{2N} & \cdots & A_{MN} \end{bmatrix} = [A_{nm}],$$

which results in an $N \times M$ matrix. If $\mathbf{A}^T = \mathbf{A}$, then $\mathbf{A}$ is said to be *symmetric* ($A_{nm} = A_{mn}$). Note that a matrix must be square to be symmetric. If instead the matrix is such that $\mathbf{A} = -\mathbf{A}^T$, it is called *skew-symmetric*. Note that for this to be true, the elements along the main diagonal of $\mathbf{A}$ must all be zero.

If the elements of $\mathbf{A}$ are complex and $\overline{\mathbf{A}}^T = \mathbf{A}$, then $\mathbf{A}$ is a *Hermitian matrix* ($\overline{A}_{nm} = A_{mn}$), where the overbar represents the *complex conjugate*, and $\overline{\mathbf{A}}^T$ is the *conjugate transpose* of $\mathbf{A}$. Note that a symmetric matrix is a special case of a Hermitian matrix.

Zero Matrix (**0**): Matrix of all zeros.

Identity Matrix (**I**): Square matrix with ones on the main diagonal and zeros everywhere else, for example,

$$\mathbf{I}_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = [\delta_{mn}],$$

where

$$\delta_{mn} = \begin{cases} 1, & m = n \\ 0, & m \neq n \end{cases}.$$

Triangular Matrix: All elements above (left triangular) or below (right triangular) the main diagonal are zero. For example,

$$\mathbf{L} = \begin{bmatrix} A_{11} & 0 & 0 & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 & 0 \\ A_{31} & A_{32} & A_{33} & 0 & 0 \\ A_{41} & A_{42} & A_{43} & A_{44} & 0 \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ 0 & A_{22} & A_{23} & A_{24} & A_{25} \\ 0 & 0 & A_{33} & A_{34} & A_{35} \\ 0 & 0 & 0 & A_{44} & A_{45} \\ 0 & 0 & 0 & 0 & A_{55} \end{bmatrix}.$$

Tridiagonal Matrix: All elements are zero except along the lower (first subdiagonal), main, and upper (first superdiagonal) diagonals as follows:

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & 0 & 0 & 0 \\ A_{21} & A_{22} & A_{23} & 0 & 0 \\ 0 & A_{32} & A_{33} & A_{34} & 0 \\ 0 & 0 & A_{43} & A_{44} & A_{45} \\ 0 & 0 & 0 & A_{54} & A_{55} \end{bmatrix}.$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Hessenberg Matrix: All elements are zero below the lower diagonal, that is

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} \\ 0 & A_{32} & A_{33} & A_{34} & A_{35} \\ 0 & 0 & A_{43} & A_{44} & A_{45} \\ 0 & 0 & 0 & A_{54} & A_{55} \end{bmatrix}.$$

Toeplitz Matrix: Each diagonal is a constant, such that

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ A_{21} & A_{11} & A_{12} & A_{13} & A_{14} \\ A_{31} & A_{21} & A_{11} & A_{12} & A_{13} \\ A_{41} & A_{31} & A_{21} & A_{11} & A_{12} \\ A_{51} & A_{41} & A_{31} & A_{21} & A_{11} \end{bmatrix}.$$

Matrix Inverse: If a square matrix $\mathbf{A}$ is *invertible*, then its inverse $\mathbf{A}^{-1}$ is such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

Note: If $\mathbf{A}$ is the 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

then its inverse is

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix},$$

where the diagonal elements A_{11} and A_{22} have been exchanged, the off-diagonal elements A_{12} and A_{21} have each switched signs, and $|\mathbf{A}| = A_{11}A_{22} - A_{12}A_{21}$ is the *determinant* of $\mathbf{A}$ (see Section 1.4.3).

Orthogonal Matrix: An $N \times N$ square matrix $\mathbf{A}$ is *orthogonal* if

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}.$$

It follows that

$$\mathbf{A}^T = \mathbf{A}^{-1}$$

for an orthogonal matrix. Such a matrix is called orthogonal because its column (and row) vectors are mutually orthogonal (see Sections 1.2 and 2.4).

Block Matrix: A block matrix is comprised of smaller submatrices, called *blocks*. For example, suppose that matrix $\mathbf{A}$ is a 2×2 block matrix comprised of four submatrices as follows:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}.$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Table 1.1 Corresponding types of real and complex matrices.

Real	Complex
Symmetric: $\mathbf{A} = \mathbf{A}^T$	Hermitian: $\mathbf{A} = \overline{\mathbf{A}}^T$
Skew-symmetric: $\mathbf{A} = -\mathbf{A}^T$	Skew-Hermitian: $\mathbf{A} = -\overline{\mathbf{A}}^T$
Orthogonal: $\mathbf{A}^{-1} = \mathbf{A}^T$	Unitary: $\mathbf{A}^{-1} = \overline{\mathbf{A}}^T$

Although each submatrix need not be of the same size, blocks in the same row must have the same number of rows, and those in the same column must have the same number of columns. That is, $\mathbf{A}_{11}$ and $\mathbf{A}_{12}$ must have the same number of rows, $\mathbf{A}_{21}$ and $\mathbf{A}_{22}$ must have the same number of rows, $\mathbf{A}_{11}$ and $\mathbf{A}_{21}$ must have the same number of columns, and $\mathbf{A}_{12}$ and $\mathbf{A}_{22}$ must also have the same number of columns.

Generalizations of some of the preceding definitions to matrices with complex elements are given in Table 1.1.

1.3 Algebraic Operations

This being linear algebra, we must extend the algebraic operations of addition and multiplication, which are so familiar for scalars, to vectors and matrices.

Addition: For $\mathbf{A}$ and $\mathbf{B}$ having the same size $M \times N$, their sum is an $M \times N$ matrix obtained by adding the corresponding elements in $\mathbf{A}$ and $\mathbf{B}$, that is, $\mathbf{A} + \mathbf{B} = [A_{mn} + B_{mn}]$. For example:

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} = \begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} & A_{13}+B_{13} \\ A_{21}+B_{21} & A_{22}+B_{22} & A_{23}+B_{23} \\ A_{31}+B_{31} & A_{32}+B_{32} & A_{33}+B_{33} \end{bmatrix}.$$

Subtraction is simply a special case of addition.

Multiplication: In the case of multiplication, we must consider multiplication of a scalar and a matrix as well as a matrix and a matrix.

1. *Multiplication by a scalar*: When multiplying a scalar by a matrix, we simply multiply each element of the matrix by the scalar c. That is,

$$c\mathbf{A} = [c\, A_{mn}]\,.$$

2. *Multiplication by a matrix*: The definition of matrix multiplication is motivated by its use in linear transformations (see Sections 1.1.2 and 1.8). For the matrix product $\mathbf{AB}$ to exist, the number of columns of $\mathbf{A}$ must equal the number of rows of $\mathbf{B}$. Then

$$\underset{M\times K}{\mathbf{A}}\ \underset{K\times N}{\mathbf{B}} = \underset{M\times N}{\mathbf{C}}$$

produces an $M \times N$ matrix $\mathbf{C}$. Written out, we have

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1K} \\ A_{21} & A_{22} & \cdots & A_{2K} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mK} \\ \vdots & \vdots & & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MK} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} & \cdots & B_{1N} \\ B_{21} & B_{22} & \cdots & B_{2n} & \cdots & B_{2N} \\ \vdots & \vdots & & \vdots & & \vdots \\ B_{K1} & B_{K2} & \cdots & B_{Kn} & \cdots & B_{KN} \end{bmatrix}$$

$$= \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1N} \\ C_{21} & C_{22} & \cdots & C_{2N} \\ \vdots & \vdots & C_{mn} & \vdots \\ C_{M1} & C_{M2} & \cdots & C_{MN} \end{bmatrix},$$

where

$$C_{mn} = A_{m1}B_{1n} + A_{m2}B_{2n} + \cdots + A_{mK}B_{Kn} = \sum_{k=1}^{K} A_{mk}B_{kn}.$$

That is, C_{mn} is the *inner product* of the mth row of $\mathbf{A}$ and the nth column of $\mathbf{B}$ (see Section 1.6 for details on the inner product).

Note that in general, $\mathbf{AB} \neq \mathbf{BA}$ (even if square), that is, matrix multiplication does not commute, and *premultiplying* $\mathbf{B}$ by $\mathbf{A}$ is not the same as *postmultiplying* $\mathbf{B}$ by $\mathbf{A}$. Moreover, there is no such thing as division in linear algebra; taking its place is the concept of the matrix inverse defined in the previous section and discussed further in Section 1.5.2. More precisely, division in scalar algebra is a special case of the inverse in linear (matrix) algebra.

Rules: For matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, and the identity matrix $\mathbf{I}$ of the appropriate sizes, and l and k integer:

$$\mathbf{A}\,(\mathbf{BC}) = (\mathbf{AB})\,\mathbf{C} \quad \text{(but cannot change order of } \mathbf{A}, \mathbf{B}, \text{ and } \mathbf{C}),$$

$$\mathbf{A}\,(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC},$$

$$(\mathbf{B} + \mathbf{C})\,\mathbf{A} = \mathbf{BA} + \mathbf{CA},$$

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A},$$

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T,$$

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T \quad \text{(note reverse order)},$$

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \quad \text{(if } \mathbf{A} \text{ and } \mathbf{B} \text{ invertible; note reverse order)},$$

$$\mathbf{A}^k = \mathbf{AA}\cdots\mathbf{A} \quad (k \text{ factors}),$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

$$\mathbf{A}^{-k} = \left(\mathbf{A}^{-1}\right)^k = \mathbf{A}^{-1}\mathbf{A}^{-1}\cdots\mathbf{A}^{-1} \quad (k \text{ factors}),$$

$$\mathbf{A}^k\mathbf{A}^l = \mathbf{A}^{k+l},$$

$$\left(\mathbf{A}^k\right)^l = \mathbf{A}^{kl},$$

$$\mathbf{A}^0 = \mathbf{I}.$$

Note that scalar and vector arithmetic is a special case of matrix arithmetic, as a scalar is simply a 1×1 matrix and a vector is an $N \times 1$ matrix.

We can combine vector addition and scalar multiplication to devise a very common algebraic operation in matrix methods known as a *linear combination* of vectors. Let us say that we have the M vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M$, which are each N-dimensional. A linear combination of these vectors is given by

$$\mathbf{u} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_M\mathbf{u}_M,$$

where $a_m, m = 1, \ldots M$ are scalar constants. Thus, we are taking the combination of some portion of each vector through summing to form the N-dimensional vector $\mathbf{u}$. In Section 1.7, we will regard the vectors $\mathbf{u}_m, m = 1, \ldots, M$ to be the *basis* of a *vector space* comprised of all possible vectors $\mathbf{u}$ given by all possible values of the coefficients $a_m, m = 1, \ldots M$.

1.4 Systems of Linear Algebraic Equations – Preliminaries

The primary application of vectors and matrices is in representing and solving systems of coupled linear algebraic equations. The size of the system may be as small as 2×2, for a simple dynamical system, to $N \times N$, where N is in the millions, for systems that arise from implementation of numerical methods applied to complex physical problems. In the following, we discuss the properties of such systems and methods for determining their solution.

1.4.1 Applications

There is nothing inherently physical about the way vectors and matrices are defined and manipulated. However, they provide a convenient means of representing the mathematical models that apply in a vast array of applications that span all areas of physics, engineering, computer science, and image processing, for example. Once we become comfortable with them, vectors and matrices simply become natural extensions of the basic algebra and calculus that are so inherent to mathematics. Linear algebra, then, is not so much a separate branch of applied mathematics as it is an essential part of our mathematical repertoire.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

The most common class of problems in matrix methods is in solving systems of coupled linear algebraic equations. These arise, for example, in solving for the forces in static (nonmoving) systems of discrete objects, parallel electrical circuits containing only voltage sources and resistors, and linear coordinate transformations. Perhaps the most important application is in numerical methods, which are designed to reduce continuous differential equations to – typically very large – systems of linear algebraic equations so as to make them amenable to solution via matrix methods. We address three of these applications in the following series of examples, and additional applications will be interspersed periodically in order to remind ourselves that we are developing the tools to solve real problems and also to expand our understanding of what the general mathematical constructs represent in physical and numerical contexts.

A system of linear algebraic equations may be represented in the form

$$\mathbf{Au} = \mathbf{b},$$

where $\mathbf{A}$ is a given coefficient matrix, $\mathbf{b}$ is a given right-hand-side vector, and $\mathbf{u}$ is the solution vector that is sought. Solution techniques are covered in the next section.

Example 1.1 We seek to determine the currents u in the parallel electrical circuit shown in Figure 1.2 (see Jeffrey 2002).

Solution

Recall Ohm's law for resistors

$$V = uR,$$

where V is the voltage drop across the resistor, u is the current through the resistor, and R is its resistance. In addition, we have Kirchhoff's laws:

1. The sum of the currents into each junction is equal to the sum of the currents out of each junction.
2. The sum of the voltage drops around each closed circuit is zero.

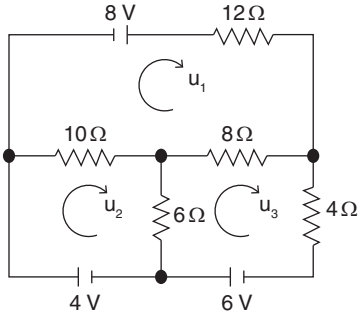


Figure 1.2 Schematic of the electrical circuit in Example 1.1.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Applying the second of Kirchhoff's laws around each of the three parallel circuits gives the three coupled equations

$$\begin{aligned} \text{Loop 1:} \quad & 8 - 12u_1 - 8(u_1 - u_3) - 10(u_1 - u_2) = 0, \\ \text{Loop 2:} \quad & 4 - 10(u_2 - u_1) - 6(u_2 - u_3) = 0, \\ \text{Loop 3:} \quad & 6 - 6(u_3 - u_2) - 8(u_3 - u_1) - 4u_3 = 0. \end{aligned}$$

Upon simplifying, this gives the system of linear algebraic equations

$$\begin{aligned} 30u_1 - 10u_2 - 8u_3 &= 8, \\ -10u_1 + 16u_2 - 6u_3 &= 4, \\ -8u_1 - 6u_2 + 18u_3 &= 6, \end{aligned}$$

which in matrix form $\mathbf{Au} = \mathbf{b}$ is

$$\begin{bmatrix} 30 & -10 & -8 \\ -10 & 16 & -6 \\ -8 & -6 & 18 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 6 \end{bmatrix}.$$

Solving this system of coupled, linear, algebraic equations would provide the sought-after solution for the currents $\mathbf{u} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T$.

Note that if the circuit also includes capacitors and/or inductors, application of Kirchhoff's laws would produce a system of ordinary differential equations as illustrated in Example 2.9.

Example 1.2 Recall that for a structure in static equilibrium, the sum of the forces and moments are both zero, which holds for each individual member as well as the entire structure. Determine the forces in each member of the truss structure shown in Figure 1.3 (see Jeffrey 2002).

Solution

Note that all members are of length ℓ; therefore, the structure is comprised of equilateral triangles with all angles being $\pi/3$. We first obtain the reaction forces at the supports 1 and 2 by summing the forces and moments for the entire structure as follows:

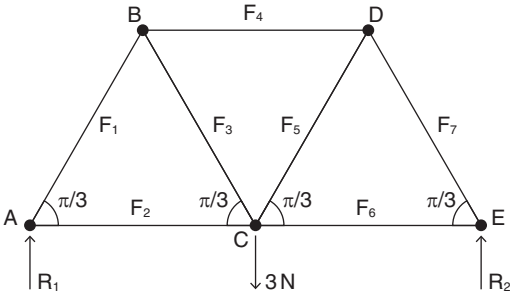


Figure 1.3 Schematic of the truss structure in Example 1.2. All members are of length ℓ.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

$$\begin{aligned} \sum F_y = 0 \quad &\Rightarrow \quad R_1 + R_2 - 3 = 0, \\ \sum M_A = 0 \quad &\Rightarrow \quad 2\ell R_2 - 3\ell = 0; \end{aligned}$$

therefore, reaction forces are

$$R_1 = \frac{3}{2}, \quad R_2 = \frac{3}{2}.$$

Next we draw a free-body diagram for each of the joints A, B, C, D, and E and sum the forces in the x- and y-directions to zero as follows:

$$\begin{aligned} \sum F_{A_x} = 0 \quad &\Rightarrow \quad F_2 + F_1 \cos\left(\tfrac{\pi}{3}\right) = 0, \\ \sum F_{A_y} = 0 \quad &\Rightarrow \quad R_1 + F_1 \sin\left(\tfrac{\pi}{3}\right) = 0, \\ \sum F_{B_x} = 0 \quad &\Rightarrow \quad F_4 + F_3 \cos\left(\tfrac{\pi}{3}\right) - F_1 \cos\left(\tfrac{\pi}{3}\right) = 0, \\ \sum F_{B_y} = 0 \quad &\Rightarrow \quad -F_1 \sin\left(\tfrac{\pi}{3}\right) - F_3 \sin\left(\tfrac{\pi}{3}\right) = 0, \\ &\vdots \end{aligned}$$

and similarly for joints C, D, and E. Because $\cos\left(\frac{\pi}{3}\right) = \frac{1}{2}$ and $\sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}$, the resulting system of equations in matrix form is

$$\begin{bmatrix} \frac{1}{2} & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & -1 & 0 & 0 & 0 \\ \frac{\sqrt{3}}{2} & 0 & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 & -\frac{1}{2} & -1 & 0 \\ 0 & 0 & \frac{\sqrt{3}}{2} & 0 & \frac{\sqrt{3}}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{\sqrt{3}}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \\ F_7 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{3}{2} \\ 0 \\ 0 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \\ -\frac{3}{2} \end{bmatrix}.$$

Carefully examining the coefficient matrix and right-hand-side vector, one can see that the coefficient matrix contains information that only relates to the structure's geometry, while the right-hand-side vector is determined by the external loads only. Hence, different loading scenarios for the same structure can be considered by simply adjusting **b** accordingly. Observe that there are ten equations (two for each joint), but only seven unknown forces (one for each member); therefore, three of the equations must be linear combinations of the other equations in order to have a unique solution (see Section 1.6).

The previous two examples illustrate how matrices can be used to represent the governing equations of a physical system directly. There are many such examples in science and engineering. In the truss problem, for example, the equilibrium equations are applied to each *discrete* element of the structure leading to a system of linear

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

algebraic equations for the forces in each member that must be solved simultaneously. In other words, the *discretization* follows directly from the geometry of the problem. For continuous systems, such as solids, fluids, and heat transfer problems involving continuous media, the governing equations are in the form of continuous ordinary or partial differential equations for which the discretization is less obvious. This gets us into the important and expansive topic of numerical methods. Although this is beyond the scope of our considerations at this time (see Part II), let us briefly motivate the overall approach using a simple one-dimensional example.

Example 1.3 Consider the fluid flow between two infinite parallel flat plates with the upper surface moving with constant speed U and an applied pressure gradient in the x-direction. This is known as Couette flow and is shown schematically in Figure 1.4.

Solution

The one-dimensional, fully developed flow is governed by the ordinary differential equation, which is derived from the Navier–Stokes equations enforcing conservation of momentum,

$$\frac{d^2u}{dy^2} = \frac{1}{\mu}\frac{dp}{dx} = \text{constant}, \tag{1.6}$$

where $u(y)$ is the fluid velocity in the x-direction, which we seek; $p(x)$ is the specified linear pressure distribution in the x-direction (such that dp/dx is a constant); and μ is the fluid viscosity. The no-slip boundary conditions at the lower and upper surfaces, are

$$u = 0 \quad \text{at} \quad y = 0, \tag{1.7}$$

$$u = U \quad \text{at} \quad y = H, \tag{1.8}$$

respectively. Equations (1.6) through (1.8) represent the mathematical model of the physical phenomenon, which is in the form of a differential equation.

In order to discretize the continuous problem, we first divide the continuous domain $0 \leq y \leq H$ into I equal subintervals of length $\Delta y = H/I$ as shown in Figure 1.5.

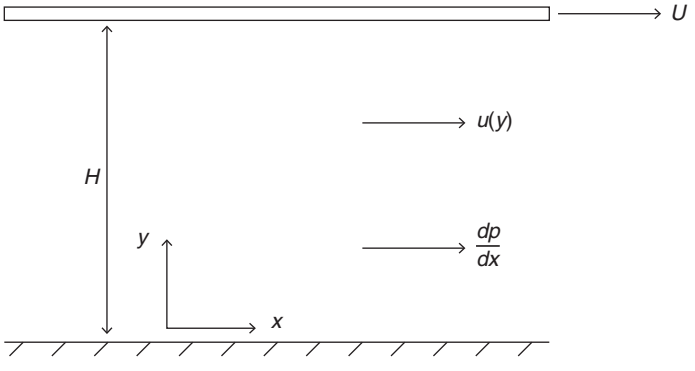


Figure 1.4 Schematic of Couette flow.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

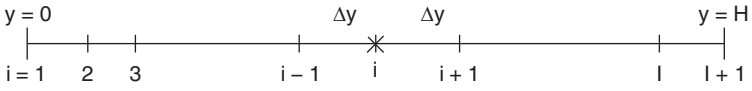


Figure 1.5 Grid used for Couette flow example. The "×" marks the location where the differential equation is approximated.

Now we must discretize the governing equation (1.6) on this *grid* or *mesh*. To see how this is done using the *finite-difference method*, recall the definition of the derivative

$$\left(\frac{du}{dy}\right)_{y=y_i} = \lim_{\Delta y \to 0} \frac{u(y_{i+1}) - u(y_i)}{\Delta y},$$

where $\Delta y = y_{i+1} - y_i$. This suggests that we can approximate the derivative by taking Δy to be small, but not in the limit as it goes all the way to zero. Thus,

$$\left(\frac{du}{dy}\right)_{y=y_i} \approx \frac{u(y_{i+1}) - u(y_i)}{\Delta y} = \frac{u_{i+1} - u_i}{\Delta y},$$

where $u_i = u(y_i)$. This is called a *forward difference*. A more accurate approximation is given by a *central difference* of the form

$$\left(\frac{du}{dy}\right)_{y=y_i} \approx \frac{u_{i+1} - u_{i-1}}{2\Delta y}. \tag{1.9}$$

We can obtain a central-difference approximation for the second-order derivative

$$\frac{d^2u}{dy^2} = \frac{d}{dy}\left(\frac{du}{dy}\right)$$

as in (1.6) by applying (1.9) midway between successive grid points as follows:

$$\left(\frac{d^2u}{dy^2}\right)_{y=y_i} \approx \frac{\left(\frac{du}{dy}\right)_{i+1/2} - \left(\frac{du}{dy}\right)_{i-1/2}}{\Delta y}.$$

Approximating du/dy as in (1.9) yields

$$\left(\frac{d^2u}{dy^2}\right)_{y=y_i} \approx \frac{\frac{u_{i+1}-u_i}{\Delta y} - \frac{u_i-u_{i-1}}{\Delta y}}{\Delta y},$$

or

$$\left(\frac{d^2u}{dy^2}\right)_{y=y_i} \approx \frac{u_{i+1} - 2u_i + u_{i-1}}{(\Delta y)^2}. \tag{1.10}$$

This is the central-difference approximation for the second-order derivative. As we will see in Chapter 7, the finite-difference approximations (1.9) and (1.10) can be formally derived using Taylor series expansions.

Given (1.10), we can discretize the governing differential (1.6) according to

$$\frac{u_{i+1} - 2u_i + u_{i-1}}{(\Delta y)^2} = \frac{1}{\mu}\frac{dp}{dx},$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

or

$$au_{i-1} + bu_i + cu_{i+1} = d, \tag{1.11}$$

where

$$a = 1, \quad b = -2, \quad c = 1, \quad d = \frac{(\Delta y)^2}{\mu}\frac{dp}{dx}.$$

The boundary conditions (1.7) and (1.8) then require that

$$u_1 = 0, \quad u_{I+1} = U. \tag{1.12}$$

Applying (1.11) at each *node point* in the grid corresponding to $i = 2,3,4,\ldots,I$ (u_1 and u_{I+1} are known) gives the system of equations

$$\begin{array}{lrcl}
i = 2: & au_1 + bu_2 + cu_3 & = & d, \\
i = 3: & au_2 + bu_3 + cu_4 & = & d, \\
i = 4: & au_3 + bu_4 + cu_5 & = & d, \\
 & & \vdots & \\
i = i: & au_{i-1} + bu_i + cu_{i+1} & = & d, \\
 & & \vdots & \\
i = I-1: & au_{I-2} + bu_{I-1} + cu_I & = & d, \\
i = I: & au_{I-1} + bu_I + cu_{I+1} & = & d.
\end{array}$$

Given the boundary conditions (1.12), the system in matrix form is

$$\begin{bmatrix}
b & c & & & & & & \\
a & b & c & & & & & \\
 & a & b & c & & & & \\
 & & \ddots & \ddots & \ddots & & & \\
 & & & a & b & c & & \\
 & & & & \ddots & \ddots & \ddots & \\
 & & & & & a & b & c \\
 & & & & & & a & b
\end{bmatrix}
\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_i \\ \vdots \\ u_{I-1} \\ u_I \end{bmatrix}
=
\begin{bmatrix} d \\ d \\ d \\ \vdots \\ d \\ \vdots \\ d \\ d - cU \end{bmatrix},$$

where all of the empty elements in the coefficient matrix are zeros. This is called a *tridiagonal Toeplitz matrix* as only three of the diagonals have nonzero elements (tridiagonal), and the elements along each diagonal are constant (Toeplitz). This structure arises because of how the equation is discretized using central differences that involve u_{i-1}, u_i, and u_{i+1}. In general, the number of domain subdivisions I can be very large, thereby leading to a large system of algebraic equations to solve for the velocities $u_2, u_3, \ldots, u_i, \ldots, u_I$.

In all three examples, we end up with a system of linear algebraic equations to solve for the currents, forces, or discretized velocities. We will encounter numerous additional examples throughout the text.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

1.4.2 General Considerations

A linear equation is one in which only polynomial terms of first degree or less appear. For example, a linear equation in N variables, $u_1, u_2, \ldots, u_N$, is of the form

$$A_1 u_1 + A_2 u_2 + \cdots + A_N u_N = b,$$

where $A_n, n = 1, \ldots, N$ and b are constants. A system of M coupled linear equations for N variables is of the form

$$\begin{aligned} &A_{11} u_1 + A_{12} u_2 + \cdots + A_{1N} u_N = b_1, \\ &A_{21} u_1 + A_{22} u_2 + \cdots + A_{2N} u_N = b_2, \\ &\vdots \\ &A_{M1} u_1 + A_{M2} u_2 + \cdots + A_{MN} u_N = b_M. \end{aligned}$$

The solution vector $\mathbf{u} = \begin{bmatrix} u_1 & u_2 & \cdots & u_N \end{bmatrix}^T$ must satisfy all M equations simultaneously. This system may be written in matrix form $\mathbf{Au} = \mathbf{b}$ as

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix}.$$

Note that the coefficient matrix $\mathbf{A}$ is $M \times N$, the solution vector $\mathbf{u}$ is $N \times 1$, and the right-hand-side vector $\mathbf{b}$ is $M \times 1$ according to the rules of matrix multiplication.

REMARKS:

1. *A system of equations may be considered a transformation, in which premultiplication by the matrix* $\mathbf{A}$ *transforms the vector* $\mathbf{u}$ *into the vector* $\mathbf{b}$ *(see Section 1.8).*
2. *A system with* $\mathbf{b} = \mathbf{0}$ *is said to be homogeneous.*
3. *The solution of* $\mathbf{Au} = \mathbf{b}$ *is the list of values* $\mathbf{u}$ *at which all of the equations intersect. For example, consider two lines* l_1 *and* l_2 *given by*

   $$A_{11} u_1 + A_{12} u_2 = b_1,$$

   and

   $$A_{21} u_1 + A_{22} u_2 = b_2,$$

   respectively. As shown in Figure 1.6, the possibilities are:
   1. l_1 *and* l_2 *have a unique intersection point at* $(u_1, u_2) \Rightarrow$ *one unique solution.*
   2. l_1 *and* l_2 *are the same line* $\Rightarrow$ *infinite solutions.*
   3. l_1 *and* l_2 *are parallel lines* $\Rightarrow$ *no solution (inconsistent).*

   The same three possibilities hold for systems of any size.
4. *As in Strang (2006), one can also interpret a system of linear algebraic equations as a linear combination of the column vectors of* $\mathbf{A}$*. If the column vectors of* $\mathbf{A}$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

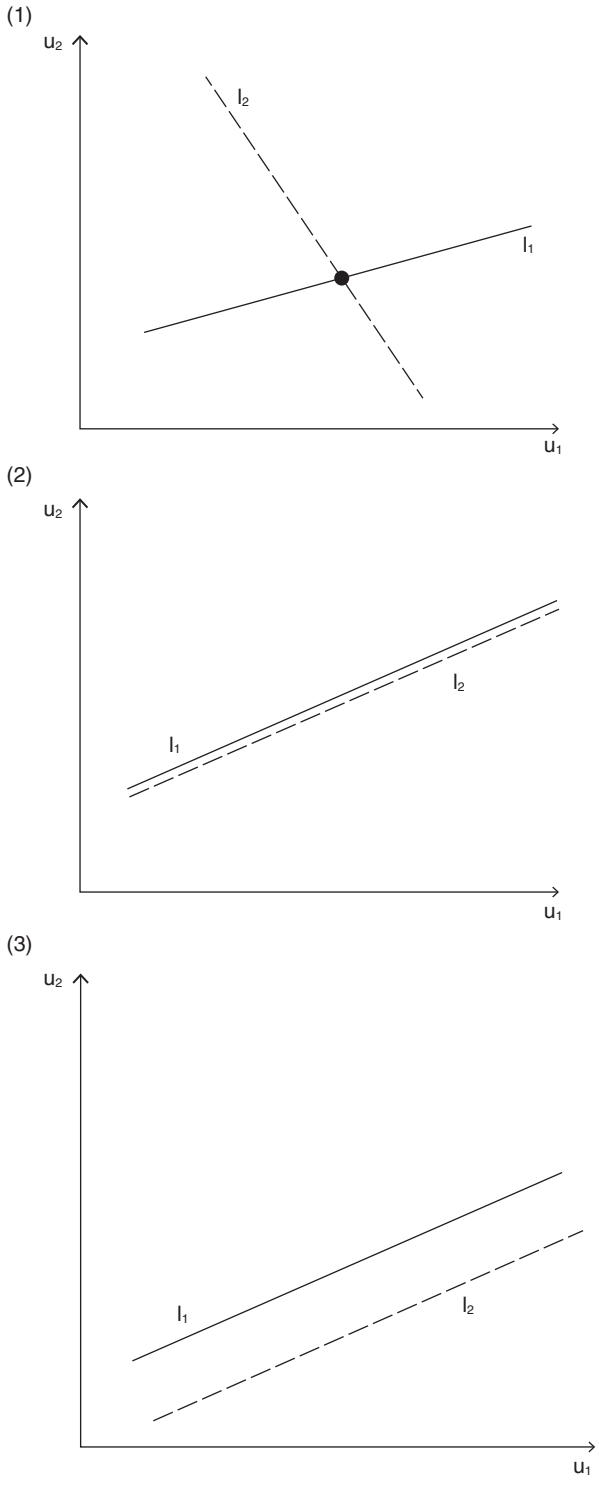


Figure 1.6 Three possibilities for a system of two linear algebraic equations.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

are denoted by $\mathbf{a}_n, n = 1, \ldots N$*, then the system of equations* $\mathbf{Au} = \mathbf{b}$ *can be represented as a linear combination as follows:*

$$u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + \cdots + u_N\mathbf{a}_N = \mathbf{b}.$$

In this interpretation, the solution vector $\mathbf{u}$ *contains the "amounts" of each of the column vectors required to form the right-hand-side vector* $\mathbf{b}$*, that is, the linear combination of the column vectors that produces* $\mathbf{b}$*. This interpretation emphasizes the fact that the right-hand-side vector must be in the column space of the matrix* $\mathbf{A}$*, that is, the vector space having the column vectors of* $\mathbf{A}$ *as its basis, in order to have a unique solution (see Section 1.7.4).*

1.4.3 Determinants

When solving systems of linear algebraic equations, the *determinant* of the coefficient matrix proves to be a useful diagnostic as well as being instrumental in solving such systems. For example, it allows us to determine whether the coefficient matrix is invertible or not and, thus, whether the system of equations has a unique solution, an infinity of solutions, or no solutions. It also appears in Cramer's rule for solving systems of algebraic equations in Section 1.5.3. We first motivate the determinant and why it is useful, then describe how to compute it for moderate-sized matrices, and finally describe some of its properties.

Consider the system of linear algebraic equations

$$\begin{aligned} A_{11}u_1 + A_{12}u_2 &= b_1, \\ A_{21}u_1 + A_{22}u_2 &= b_2, \end{aligned}$$

where A_{11}, A_{12}, A_{21}, A_{22}, b_1, and b_2 are known constants, and u_1 and u_2 are the variables to be found. In order to solve for the unknown variables, multiply the first equation by A_{22}, the second by A_{12}, and then subtract. This eliminates the u_2 variable, and solving for u_1 gives

$$u_1 = \frac{A_{22}b_1 - A_{12}b_2}{A_{11}A_{22} - A_{12}A_{21}}.$$

Similarly, to obtain u_2, multiply the first equation by A_{21}, the second by A_{11}, and then subtract. This eliminates the u_1 variable, and solving for u_2 gives

$$u_2 = \frac{A_{11}b_2 - A_{21}b_1}{A_{11}A_{22} - A_{12}A_{21}}.$$

Observe that the denominators are the same in both cases. We call this denominator the *determinant* of the 2×2 coefficient matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

and denote it by

$$|\mathbf{A}| = A_{11}A_{22} - A_{12}A_{21}.$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Note that $|\mathbf{A}| \neq 0$ for there to be a solution for u_1 and u_2. If $|\mathbf{A}| = 0$, we say that $\mathbf{A}$ is *singular*.

For a 3×3 matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \tag{1.13}$$

the determinant is given by

$$\begin{aligned} |\mathbf{A}| = {} & A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32} \\ & - A_{13}A_{22}A_{31} - A_{12}A_{21}A_{33} - A_{11}A_{23}A_{32}. \end{aligned} \tag{1.14}$$

Note that this diagonal multiplication method does not work for $\mathbf{A}$ larger than 3×3. Therefore, a *cofactor expansion* must be used.

For a square $N \times N$ matrix $\mathbf{A}$, M_{mn} is the *minor* and C_{mn} is the *cofactor* of A_{mn}. The minor M_{mn} is the determinant of the submatrix that remains after the mth row and nth column are removed, and the cofactor of $\mathbf{A}$ is

$$C_{mn} = (-1)^{m+n} M_{mn}.$$

We can then write the *cofactor matrix*, which is

$$\mathbf{C} = [C_{mn}] = \begin{bmatrix} +M_{11} & -M_{12} & +M_{13} & -M_{14} & \cdots \\ -M_{21} & +M_{22} & -M_{23} & +M_{24} & \\ +M_{31} & -M_{32} & +M_{33} & -M_{34} & \\ -M_{41} & +M_{42} & -M_{43} & +M_{44} & \\ \vdots & & & & \ddots \end{bmatrix},$$

where we note the sign pattern for the cofactors.

A *cofactor expansion* can be used to obtain determinants by converting the determinant of a large matrix into a sum of terms involving determinants of smaller matrices that can be easily evaluated. To illustrate, consider the determinant of 3×3 matrix $\mathbf{A}$ given by (1.13). The cofactor expansion can be formed using any row or column of matrix $\mathbf{A}$. For example, using the first row

$$\begin{aligned} |\mathbf{A}| &= A_{11}C_{11} + A_{12}C_{12} + A_{13}C_{13} \\ &= A_{11}[+M_{11}] + A_{12}[-M_{12}] + A_{13}[+M_{13}] \\ &= A_{11}\left[+\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix}\right] + A_{12}\left[-\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix}\right] + A_{13}\left[+\begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix}\right] \\ &= A_{11}\left(A_{22}A_{33} - A_{23}A_{32}\right) - A_{12}\left(A_{21}A_{33} - A_{23}A_{31}\right) \\ &\quad + A_{13}\left(A_{21}A_{32} - A_{22}A_{31}\right) \\ |\mathbf{A}| &= A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32} \\ &\quad - A_{11}A_{23}A_{32} - A_{12}A_{21}A_{33} - A_{13}A_{22}A_{31}, \end{aligned}$$

which is the same as (1.14).

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

The cofactor matrix is used in Cramer's rule in Section 1.5.3, and the cofactor expansion is used to obtain the cross product of two vectors in Section 1.6.1.

Properties of Determinants ($\mathbf{A}$ and $\mathbf{B}$ are $N \times N$):

1. If any row or column of $\mathbf{A}$ is all zeros, $|\mathbf{A}| = 0$.
2. If two rows (or columns) of $\mathbf{A}$ are interchanged, the sign of $|\mathbf{A}|$ changes.
3. If any row (or column) of $\mathbf{A}$ is a linear combination of the other rows (or columns), then $|\mathbf{A}| = 0$.
4. $|\mathbf{A}^T| = |\mathbf{A}|$.
5. In general, $|\mathbf{A} + \mathbf{B}| \neq |\mathbf{A}| + |\mathbf{B}|$.
6. $|\mathbf{AB}| = |\mathbf{BA}| = |\mathbf{A}||\mathbf{B}|$.
7. If $|\mathbf{A}| = 0$, then $\mathbf{A}$ is *singular* and not invertible. If $|\mathbf{A}| \neq 0$, then $\mathbf{A}$ is *nonsingular* and invertible, and $|\mathbf{A}^{-1}| = \dfrac{1}{|\mathbf{A}|}$.
8. The determinant of a triangular matrix is the product of the main diagonal elements.
9. The determinant of an orthogonal matrix is $|\mathbf{A}| = \pm 1$ (see Section 2.4).

1.5 Systems of Linear Algebraic Equations – Solution Methods

Let us turn our attention to basic methods for solving systems of linear algebraic equations. Methods developed in this chapter are suitable for hand calculations involving small systems of equations or using mathematics software, such as MATLAB or Mathematica,[7] for moderate-sized systems. Techniques for solving very large systems of equations computationally are discussed in Part II, and all of the methods discussed in this text for solving systems of linear algebraic equations will be summarized in Table 6.1 with reference to their respective sections.

1.5.1 Gaussian Elimination

Although not always the most efficient means of solution, *Gaussian elimination* is the most general approach, the starting point for certain specialized techniques, and the standard by which other methods are compared. Its primary virtue is that it can be applied to *any* system of equations to determine its solution(s), and indeed decipher whether it even has a solution.

The strategy in Gaussian elimination is to solve the system of linear equations $\mathbf{Au} = \mathbf{b}$ by first converting it to a new system that has the same solution $\mathbf{u}$ but is straightforward to solve. This is accomplished using *elementary row operations*, which do not change the solution vector, in a prescribed sequence. The elementary row operations correspond to actions that could be applied to a set of equations without changing their solution and are:

[7] MATLAB and Mathematica are commercial products of MathWorks and Wolfram Research, respectively.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

1. Interchange two equations (rows).
2. Multiply an equation (row) by a nonzero number.
3. Add a multiple of one equation (row) to another.

But what is the objective of the elementary row operations? Observe that if our system were such that the coefficient matrix was simply an identity matrix, then the equations would be uncoupled. In this case, the solution of each equation can be written down by inspection. This is known as the *reduced row-echelon form* of the system of equations. Requiring far fewer elementary row operations, and only a bit more work to obtain the solution, we seek the *row-echelon form* of the system of equations. This consists of ones down the main diagonal with zeros below. Therefore, it is a right, or upper, triangular matrix. Thus, the objective of Gaussian elimination is to perform a sequence of elementary row operations that zero out the elements below the main diagonal and leaves leading ones in each row.

Procedure:

1. Write the *augmented matrix*:

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{cccc|c} A_{11} & A_{12} & \cdots & A_{1N} & b_1 \\ A_{21} & A_{22} & \cdots & A_{2N} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} & b_M \end{array}\right].$$

2. Move rows (equations) with leading zeros to the bottom.
3. Divide the first row by A_{11}:

$$\left[\begin{array}{cccc|c} 1 & \hat{A}_{12} & \cdots & \hat{A}_{1N} & \hat{b}_1 \\ A_{21} & A_{22} & \cdots & A_{2N} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} & b_M \end{array}\right].$$

4. Multiply the first row by A_{m1} and subtract from the mth row, where $m = 2, 3, \ldots, M$:

$$\left[\begin{array}{cccc|c} 1 & \hat{A}_{12} & \cdots & \hat{A}_{1N} & \hat{b}_1 \\ 0 & \hat{A}_{22} & \cdots & \hat{A}_{2N} & \hat{b}_2 \\ 0 & \hat{A}_{32} & \cdots & \hat{A}_{3N} & \hat{b}_3 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \hat{A}_{M2} & \cdots & \hat{A}_{MN} & \hat{b}_M \end{array}\right].$$

5. Repeat steps (2) – (4), called *forward elimination*, on the lower-right $(M-1) \times (N-1)$ submatrix until the matrix is in row-echelon form, in which case the first nonzero element in each row is a one with zeros below it.
6. Obtain the solution by *backward substitution*. First, solve each equation for its leading variable (for example, $u_1 = \ldots, u_2 = \ldots, u_3 = \cdots$). Second, starting at the bottom, back substitute each equation into all of the equations above it. Finally,

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

the remaining $N - \text{rank}(\mathbf{A})$ variables are arbitrary, where the *rank* of matrix $\mathbf{A}$, denoted by $\text{rank}(\mathbf{A})$, is defined below.

Example 1.4 Using Gaussian elimination, solve the system of linear algebraic equations

$$\begin{aligned} 2u_1 + u_3 &= 3, \\ u_2 + 2u_3 &= 3, \\ u_1 + 2u_2 &= 3. \end{aligned}$$

Solution

In the matrix form $\mathbf{Au} = \mathbf{b}$, this system of equations is represented in the form

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

Note that the determinant of $\mathbf{A}$ is $|\mathbf{A}| = -9$; therefore, the matrix is nonsingular, and we expect a unique solution.

In order to apply Gaussian elimination, let us write the augmented matrix

$$\left[\begin{array}{ccc|c} 2 & 0 & 1 & 3 \\ 0 & 1 & 2 & 3 \\ 1 & 2 & 0 & 3 \end{array}\right],$$

and perform elementary row operations to produce the row-echelon form via forward elimination. In order to have a leading one in the first column of the first row, dividing through by two produces

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 1 & 2 & 3 \\ 1 & 2 & 0 & 3 \end{array}\right].$$

We first seek to eliminate the elements below the leading one in the first column of the first row. The first column of the second row already contains a zero, so we focus on the third row. To eliminate the leading one, subtract the third row from the first row to produce

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 1 & 2 & 3 \\ 0 & -2 & \frac{1}{2} & -\frac{3}{2} \end{array}\right].$$

Now the submatrix that results from eliminating the first row and first column is considered. There is already a leading one in the second column of the second row, so we seek to eliminate the element directly beneath it. To do so, multiply the second row by two and add to the third row to yield

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 1 & 2 & 3 \\ 0 & 0 & \frac{9}{2} & \frac{9}{2} \end{array}\right].$$

To obtain row-echelon form, divide the third row by 9/2 to give

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

Having obtained the row-echelon form, we now perform the backward substitution to obtain the solution for **u**. Beginning with the third row, we see that

$$u_3 = 1.$$

Then from the second row

$$u_2 + 2u_3 = 3,$$

or substituting $u_3 = 1$, we have

$$u_2 = 1.$$

Similarly, from the first row,

$$u_1 + \frac{1}{2}u_3 = \frac{3}{2},$$

or again substituting $u_3 = 1$ leads to

$$u_1 = 1.$$

Therefore, the solution is

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Substitution will reveal that this solution satisfies all three of the original equations.

The preceding example consisted of three equations for the three unknowns, and a unique solution was obtained through Gaussian elimination. Recall, however, that a system of equations also may produce an infinity of solutions or no solution at all. Once again, Gaussian elimination will reveal either of these three possibilities. However, there are additional diagnostics that will help determine the outcome based on evaluating the ranks of the coefficient and augmented matrices and comparing with the number of unknowns.

The *rank* of matrix **A**, denoted by rank(**A**), is given by the size of the largest square submatrix of **A** with a nonzero determinant. Equivalently, the rank of **A** is the number of nonzero rows when reduced to row-echelon form $\hat{\mathbf{A}}$. This is because elementary row operations do not change the rank of a matrix. If the rank of **A** and the rank of the augmented matrix [**A**|**b**], denoted by rank($\mathbf{A}_{aug}$), are equal, the system is said to

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

be *consistent* and the system has one or more solutions. If they are not the same, then the system is *inconsistent*, and it has no solutions.

With M being the number of equations (rows), N being the number of unknowns (columns), rank($\mathbf{A}$) being the rank of the coefficient matrix $\mathbf{A}$, and rank($\mathbf{A}_{aug}$) being the rank of the augmented matrix $\mathbf{A}_{aug}$, the three possibilities are as follows:

1. If rank($\mathbf{A}$) = rank($\mathbf{A}_{aug}$) = N, the system is consistent with as many independent equations as unknowns, such that there is a unique solution.
2. If rank($\mathbf{A}$) = rank($\mathbf{A}_{aug}$) < N, the system is consistent but with fewer independent equations than unknowns. Then there is an [N − rank($\mathbf{A}$)]-parameter family of solutions, such that [N −rank($\mathbf{A}$)] of the variables are arbitrary and there are *infinite solutions*. Because there are fewer equations than required for a unique solution, the system is said to be *underdetermined*. In such cases, the *least-squares method* can be employed to obtain the solution that satisfies an additional criterion (see Section 10.2.2).
3. If rank($\mathbf{A}$) < rank($\mathbf{A}_{aug}$),[8] then the $\hat{A}_{ij}$ elements for the ith row [i > rank($\mathbf{A}$)] are all zero, but at least one of the corresponding $\hat{b}_i$ is nonzero, and the system is *inconsistent* and *no solution* exists. Because there are more equations than required for a unique solution, the system is said to be *overdetermined*. In such cases, the *least-squares method* can be employed to obtain the solution that is "closest" to satisfying the equations (see Section 10.2.1).

Next we consider a series of three closely related examples that illustrate each of the three possible outcomes when seeking a solution of a system of linear algebraic equations as summarized in Figure 1.6.

Example 1.5 Using Gaussian elimination, solve the system of linear algebraic equations

$$\begin{aligned} u_1 + 2u_2 + 3u_3 &= 1, \\ 2u_1 + 3u_2 + 4u_3 &= 2, \\ 3u_1 + 4u_2 + 5u_3 &= 3. \end{aligned}$$

Solution

In the matrix form $\mathbf{Au} = \mathbf{b}$, this system of equations is represented in the form

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Despite the clear pattern in the coefficient matrix and right-hand-side vector, there is nothing particularly special about this matrix problem. Note that the determinant of $\mathbf{A}$ is $|\mathbf{A}| = 0$; therefore, the matrix is singular, in which case we either have no solution or an infinite number of solutions. That is, there is no unique solution.

[8] Note that it is not possible for rank($\mathbf{A}$) > rank($\mathbf{A}_{aug}$).

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

In order to apply Gaussian elimination, let us write the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 2 & 3 & 4 & 2 \\ 3 & 4 & 5 & 3 \end{array}\right],$$

and perform elementary row operations to produce the row-echelon form via forward elimination. Because there is already a leading one in the first column of the first row, we proceed to eliminating the elements in rows two and three below the leading one in the first row. Multiplying two times the first row and subtracting the second row, as well as multiplying three times the first row and subtracting the third row yields

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right].$$

There is already a leading one in the second row, so we proceed to eliminate the element below it in the third row. To do so, multiply two times the second row and subtract the third row as follows:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

which is in row-echelon form.

Observe that while we have three unknowns ($N = 3$), the rank of the coefficient matrix $\mathbf{A}$ is $\text{rank}(\mathbf{A}) = 2$. Recall that the rank of the coefficient matrix is equal to the number of nonzero rows after row reduction. The rank of the augmented matrix is $\text{rank}(\mathbf{A}_{aug}) = 2$ as well. Because the ranks of the coefficient and augmented matrices are the same, the system of equations is consistent, but because $\text{rank}(\mathbf{A}) < N$, we have an $[N - \text{rank}(\mathbf{A})]$-parameter family of solutions. That is, there are an infinite number of solutions. Specifically, we have the two equations for the three unknowns

$$\begin{aligned} u_1 + 2u_2 + 3u_3 &= 1, \\ u_2 + 2u_3 &= 0. \end{aligned}$$

The second equation requires that $u_2 = -2u_3$, which when substituted into the first equation leads to $u_1 = u_3 + 1$. Because u_3 is left to be arbitrary, we have a one-parameter family (infinity) of solutions.

In the next example, we alter the right-hand-side vector $\mathbf{b}$ from the previous example but leave the coefficient matrix $\mathbf{A}$ unaltered.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Example 1.6 Using Gaussian elimination, solve the system of linear algebraic equations

$$\begin{aligned} u_1 + 2u_2 + 3u_3 &= 3, \\ 2u_1 + 3u_2 + 4u_3 &= 1, \\ 3u_1 + 4u_2 + 5u_3 &= 2. \end{aligned}$$

Solution

In the matrix form $\mathbf{Au} = \mathbf{b}$, this system of equations is represented in the form

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}.$$

Because the coefficient matrix is unchanged from the previous example, the determinant is still zero, and the same series of elementary row operations leads to the row-echelon form.

Writing the augmented matrix gives

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 5 & 2 \end{array}\right].$$

Multiplying two times the first row and subtracting the second row, as well as multiplying three times the first row and subtracting the third row, yields

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 1 & 2 & 5 \\ 0 & 2 & 4 & 7 \end{array}\right].$$

Multiplying two times the second row and subtracting the third row yields the row-echelon form

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 0 & 3 \end{array}\right].$$

Note that while the rank of the coefficient matrix is still $\text{rank}(\mathbf{A}) = 2$, the rank of the augmented matrix is now $\text{rank}(\mathbf{A}_{aug}) = 3$. Because the ranks of the coefficient and augmented matrices are no longer the same, the system of equations is inconsistent, and there is no solution. Observe that the third equation requires that

$$0u_1 + 0u_2 + 0u_3 = 3,$$

which of course is not satisfied by any solution $\mathbf{u}$. The fact that it is not possible to satisfy this equation reflects the inconsistency of the system of equations.

Again, let us change the right-hand-side vector $\mathbf{b}$ and also one element in the coefficient matrix $\mathbf{A}$.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Example 1.7 Using Gaussian elimination, solve the system of linear algebraic equations

$$\begin{aligned} u_1 + 2u_2 + 3u_3 &= 1, \\ 2u_1 + 3u_2 + 4u_3 &= 1, \\ 3u_1 + 4u_2 + 6u_3 &= 1. \end{aligned}$$

Solution

In the matrix form $\mathbf{Au} = \mathbf{b}$, this system of equations is represented in the form

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

The determinant of $\mathbf{A}$ is now $|\mathbf{A}| = -1$; therefore, the matrix is nonsingular and invertible, and the system of equations has a unique solution. Again, we perform the same sequence of elementary row operations as only the last element has been changed in the coefficient matrix.

Writing the augmented matrix gives

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 6 & 1 \end{array}\right].$$

Multiplying two times the first row and subtracting the second row, as well as multiplying three times the first row and subtracting the third row, yields

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 2 & 3 & 2 \end{array}\right].$$

Multiplying two times the second row and subtracting the third row yields the row-echelon form

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

In this case, the ranks of the coefficient and augmented matrices are the same and equal the number of unknowns, such that $N = \text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}_{aug}) = 3$, which reflects the fact that the system is consistent and has a unique solution.

This unique solution is obtained using backward substitution as follows. The third equation indicates that

$$u_3 = 0,$$

which when substituted into the second equation requires that

$$u_2 = 1.$$

Finally, substituting both of these results into the first equation yields

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

$$u_1 = -1.$$

Therefore, the solution vector is

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Recall that if $\mathbf{A}$ is an $N \times N$ triangular matrix, then

$$|\mathbf{A}| = A_{11}A_{22}\cdots A_{NN},$$

in which case the determinant is the product of the elements on the main diagonal. Therefore, we may row reduce the matrix $\mathbf{A}$ to triangular form in order to simplify the determinant calculation. In doing so, however, we must take into account the influence of each elementary row operation on the determinant as follows:

- Interchange rows $\Rightarrow |\hat{\mathbf{A}}| = -|\mathbf{A}|$.
- Multiply the row by k $\Rightarrow |\hat{\mathbf{A}}| = k|\mathbf{A}|$.
- Add a multiple of one row to another $\Rightarrow |\hat{\mathbf{A}}| = |\mathbf{A}|$.

1.5.2 Matrix Inverse

Given a system of equations $\mathbf{Au} = \mathbf{b}$, with an invertible coefficient matrix, having $|\mathbf{A}| \neq 0$, an alternative to Gaussian elimination for solving the system is based on using the matrix inverse. Premultiplying by $\mathbf{A}^{-1}$ gives

$$\mathbf{A}^{-1}\mathbf{Au} = \mathbf{A}^{-1}\mathbf{b}.$$

Because $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, the solution vector may be obtained from

$$\mathbf{u} = \mathbf{A}^{-1}\mathbf{b}.$$

Note that one can easily change $\mathbf{b}$ and premultiply by $\mathbf{A}^{-1}$ to obtain a solution for a different right-hand-side vector. If $\mathbf{b} = \mathbf{0}$, and the system is *homogeneous*, the only solution is the trivial solution $\mathbf{u} = \mathbf{0}$.[9]

One technique for obtaining the matrix inverse is to use Gaussian elimination. This is based on the fact that the elementary row operations that reduce $\mathbf{A}_N$ to $\mathbf{I}_N$ also reduce $\mathbf{I}_N$ to $\mathbf{A}_N^{-1}$. Therefore, we augment the coefficient and identity matrices according to

$$\left[\ \mathbf{A}_N \ \middle|\ \mathbf{I}_N\ \right],$$

and perform the elementary row operations that reduce this to

$$\left[\ \mathbf{I}_N \ \middle|\ \mathbf{A}_N^{-1}\ \right].$$

Note that $\mathbf{A}$ is not invertible if the row operations do not produce $\mathbf{I}_N$.

[9] Note that while the terms *homogeneous* and *trivial* both have to do with something being zero, homogeneous refers to the right-hand side of an equation, while trivial refers to the solution.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Example 1.8 As in Example 1.4 using Gaussian elimination, consider the system of linear algebraic equations

$$\begin{aligned} 2u_1 + u_3 &= 3, \\ u_2 + 2u_3 &= 3, \\ u_1 + 2u_2 &= 3. \end{aligned}$$

Solution

In matrix form $\mathbf{Au} = \mathbf{b}$, this is

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

Recall that the determinant of $\mathbf{A}$ is $|\mathbf{A}| = -9$; therefore, the matrix is nonsingular and invertible.

Augmenting the coefficient matrix $\mathbf{A}$ with the 3×3 identify matrix $\mathbf{I}$ gives

$$\left[\begin{array}{ccc|ccc} 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

Performing the same Gaussian elimination steps as in Example 1.4 to obtain the row-echelon form produces

$$\left[\begin{array}{ccc|ccc} 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{9} & \frac{4}{9} & -\frac{2}{9} \end{array}\right].$$

To obtain the inverse of $\mathbf{A}$ on the right, we must have the identify matrix on the left, that is, we must have the left side in reduced row-echelon form. To eliminate the two in the third column of the second row, multiply the third row by -2 and add to the second row. Similarly, to eliminate the one-half in the third column of the first row, multiply the third row by $-1/2$ and add to the first row. This leads to

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{4}{9} & -\frac{2}{9} & \frac{1}{9} \\ 0 & 1 & 0 & -\frac{2}{9} & \frac{1}{9} & \frac{4}{9} \\ 0 & 0 & 1 & \frac{1}{9} & \frac{4}{9} & -\frac{2}{9} \end{array}\right].$$

Therefore, the inverse of $\mathbf{A}$ is

$$\mathbf{A}^{-1} = \frac{1}{9} \begin{bmatrix} 4 & -2 & 1 \\ -2 & 1 & 4 \\ 1 & 4 & -2 \end{bmatrix},$$

and the solution vector can then be obtained as follows:

$$\mathbf{u} = \mathbf{A}^{-1}\mathbf{b} = \frac{1}{9} \begin{bmatrix} 4 & -2 & 1 \\ -2 & 1 & 4 \\ 1 & 4 & -2 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 9 \\ 9 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

which is the same solution as obtained using Gaussian elimination directly in Example 1.4.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

See Chapter 6 for additional methods for determining the inverse of a matrix that are suitable for computer algorithms, including LU decomposition and Cholesky decomposition.

1.5.3 Cramer's Rule

Cramer's rule is an alternative method for obtaining the unique solution of a system $\mathbf{Au} = \mathbf{b}$, where $\mathbf{A}$ is $N \times N$ and nonsingular ($|\mathbf{A}| \neq 0$). Recall that the cofactor matrix for an $N \times N$ matrix $\mathbf{A}$ is

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1N} \\ C_{21} & C_{22} & \cdots & C_{2N} \\ \vdots & \vdots & & \vdots \\ C_{N1} & C_{N2} & \cdots & C_{NN} \end{bmatrix},$$

where C_{mn} is the cofactor of A_{mn}. We define the *adjugate* of $\mathbf{A}$ to be the transpose of the cofactor matrix as follows:

$$\text{Adj}(\mathbf{A}) = \mathbf{C}^T.$$

The adjugate matrix is often referred to as the *adjoint* matrix; however, this can lead to confusion with the term *adjoint* used in other settings.

Evaluating the product of matrix $\mathbf{A}$ with its adjugate, it can be shown using the rules of cofactor expansions that

$$\mathbf{A}\,\text{Adj}(\mathbf{A}) = |\mathbf{A}|\,\mathbf{I}.$$

Therefore, because $\mathbf{AA}^{-1} = \mathbf{I}$ (if $\mathbf{A}$ is invertible), we can write

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\text{Adj}(\mathbf{A}).$$

This leads, for example, to the familiar result for the inverse of a 2×2 matrix given in Section 1.2.

More generally, this result leads to Cramer's rule for the solution of the system $\mathbf{Au} = \mathbf{b}$ by recognizing that the nth element of the vector $\text{Adj}(\mathbf{A})\mathbf{b}$ is equal to $|\mathbf{A}_n|$, where $\mathbf{A}_n$, $n = 1, 2, \ldots, N$, is obtained by replacing the nth column of $\mathbf{A}$ by the right-hand-side vector

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

Then the system has a unique solution, which is

$$u_1 = \frac{|\mathbf{A}_1|}{|\mathbf{A}|}, \quad u_2 = \frac{|\mathbf{A}_2|}{|\mathbf{A}|}, \quad \ldots, \quad u_N = \frac{|\mathbf{A}_N|}{|\mathbf{A}|}.$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

REMARKS:

1. *If the system is homogeneous, in which case* $\mathbf{b} = \mathbf{0}$*, then* $|\mathbf{A}_j| = 0$*, and the unique solution is the trivial solution* $\mathbf{u} = \mathbf{0}$ *as expected.*
2. *Although Cramer's rule applies for any size nonsingular matrix, it is only efficient for small systems* ($N \leq 3$) *owing to the ease of calculating determinants of* 3×3 *and smaller systems. For large systems, using Gaussian elimination is typically more practical.*

Example 1.9 Let us once again consider the system of linear algebraic equations from Examples 1.4 and 1.8, which is

$$\begin{aligned} 2u_1 + u_3 &= 3, \\ u_2 + 2u_3 &= 3, \\ u_1 + 2u_2 &= 3. \end{aligned}$$

Solution

In matrix form $\mathbf{Au} = \mathbf{b}$, the system is

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

Again, the determinant of $\mathbf{A}$ is $|\mathbf{A}| = -9$. By Cramer's rule, the solution is

$$u_1 = \frac{|\mathbf{A}_1|}{|\mathbf{A}|} = \frac{1}{-9} \begin{vmatrix} 3 & 0 & 1 \\ 3 & 1 & 2 \\ 3 & 2 & 0 \end{vmatrix} = \frac{-9}{-9} = 1,$$

$$u_2 = \frac{|\mathbf{A}_2|}{|\mathbf{A}|} = \frac{1}{-9} \begin{vmatrix} 2 & 3 & 1 \\ 0 & 3 & 2 \\ 1 & 3 & 0 \end{vmatrix} = \frac{-9}{-9} = 1,$$

$$u_3 = \frac{|\mathbf{A}_3|}{|\mathbf{A}|} = \frac{1}{-9} \begin{vmatrix} 2 & 0 & 3 \\ 0 & 1 & 3 \\ 1 & 2 & 3 \end{vmatrix} = \frac{-9}{-9} = 1,$$

which is the same as obtained using the inverse and Gaussian elimination in the earlier examples.

1.6 Vector Operations

Vectors, which are simply $N \times 1$ matrices, can serve many purposes. In mechanics, vectors represent the magnitude and direction of quantities, such as forces, moments, and velocities; in electromagnetics, they represent the magnitude and direction of the

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

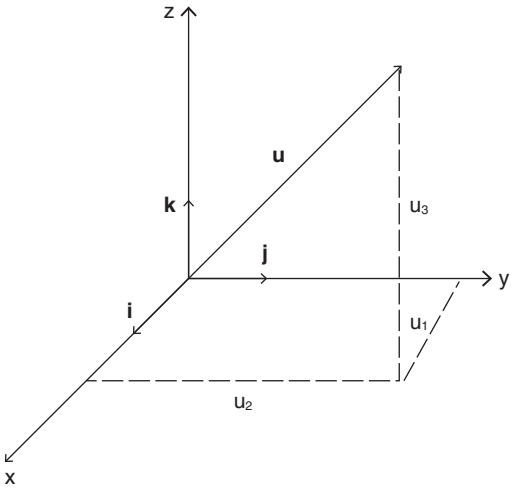


Figure 1.7 Three-dimensional vector **u** in Cartesian coordinates.

electric and magnetic fields. We first describe some operations and properties unique to vectors followed by a brief discussion of the use of vectors in forming bases for vector spaces in the next section. In order to appeal to straightforward geometric interpretations, let us first review vectors in three dimensions before extending to higher dimensions.

1.6.1 Three-Dimensional Vectors

Recall that a *scalar* is a quantity specified only by a magnitude, such as mass, length, temperature, pressure, energy, et cetera, whereas a *vector* is a quantity specified by both a magnitude and direction, such as position, velocity, force, moment, momentum, et cetera. In three dimensions, a vector **u** is given by

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix},$$

which is shown for Cartesian coordinates in Figure 1.7, in which all three basis vectors are mutually orthogonal. The unit vectors **i**, **j**, and **k** are in each of the coordinate directions. Finally, a *tensor* is a quantity specified by multiple magnitudes and directions. In three-dimensional applications, it is given by nine quantities expressed as a 3×3 matrix, such as stress, strain, moment of inertia, et cetera. See Section 4.2 for more on tensors.

Vector addition is simply a special case of matrix addition. For example, adding the two three-dimensional vectors **u** and **v** leads to

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix}.$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Observe that adding two vectors results in another vector called the *resultant vector*.

Multiplying a scalar times a vector simply scales the length of the vector accordingly without changing its direction. It is a special case of matrix-scalar multiplication. For example,

$$k\mathbf{u} = k\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} ku_1 \\ ku_2 \\ ku_3 \end{bmatrix},$$

which is a vector.

Matrix-matrix multiplication applied to vectors is known as the *inner product*.[10] The inner product of $\mathbf{u}$ and $\mathbf{v}$ is denoted by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T\mathbf{v} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = u_1v_1 + u_2v_2 + u_3v_3 = \mathbf{v}^T\mathbf{u} = \langle \mathbf{v}, \mathbf{u} \rangle,$$

which is a scalar.[11] If $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, we say that the vectors $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal*. In two and three dimensions, this means that they are geometrically perpendicular.

The inner product can be used to determine the length of a vector, which is known as its *norm*, and given by

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = \sqrt{u_1^2 + u_2^2 + u_3^2}.$$

A *unit vector* is one having length (norm) equal to one. The inner product and norm can also be used to determine the angle between two vectors. From the Pythagoras theorem, the angle θ between two vectors, as illustrated in Figure 1.8, is such that

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}. \tag{1.15}$$

If $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, then $\theta = \frac{\pi}{2}$, and the vectors are orthogonal. For more on vector norms, see Section 1.9.

A vector operation that is unique to the two- and three-dimensional cases is the *cross product*.[12] Recall that the inner (dot or scalar) product of two vectors produces a scalar; the cross product of two vectors produces a vector.

In Cartesian coordinates, where the vectors $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$ and $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$, the cross product is

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

Note how the cofactor expansion about the first row is used to obtain each component of the cross product. The cross product produces a vector $\mathbf{w}$ that is perpendicular to the plane in which the vectors $\mathbf{u}$ and $\mathbf{v}$ lie and points in the direction according to

[10] The inner product also goes by the terms *dot product* and *scalar product*.

[11] Different authors use various notation for the inner product, including $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = (\mathbf{u}, \mathbf{v})$.

[12] The cross product is sometimes referred to as a *vector product* because it produces a vector in contrast to the inner (scalar) product, which produces a scalar.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

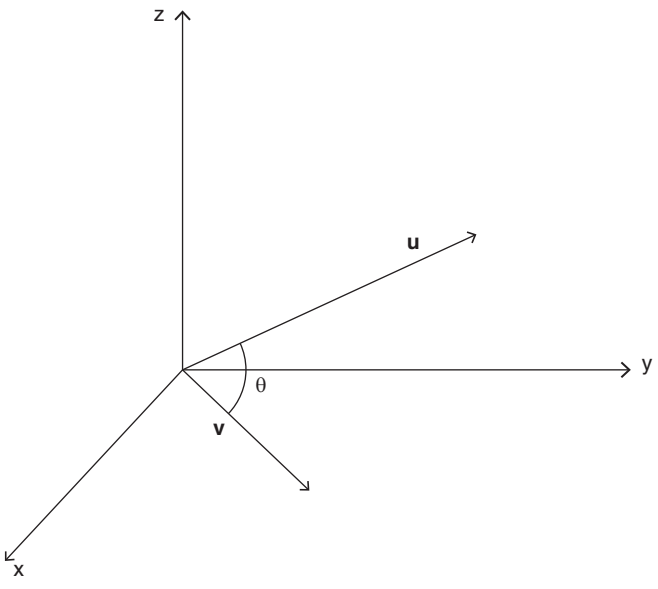


Figure 1.8 Angle between two vectors.

the right-hand rule in sweeping from $\mathbf{u}$ to $\mathbf{v}$. See Section 4.1 for the cross product in cylindrical and spherical coordinates.

REMARKS:

1. *Observe that the order of the vectors matters when taking the cross product according to the right-hand rule, whereas it does not for the inner product. That is, while* $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ *produces a scalar of the same magnitude,* $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$*, that is, the two cross products each produce a vector of the same magnitude but of opposite sign.*
2. *If either* $\mathbf{u}$ *or* $\mathbf{v}$ *is zero, or if* $\mathbf{u}$ *and* $\mathbf{v}$ *are parallel, then the cross product is the zero vector* $\mathbf{w} = \mathbf{0}$*.*
3. *If the two vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *are taken to be the two adjacent sides of a parallelogram, then the norm (length) of the cross product* $\|\mathbf{w}\| = \|\mathbf{u} \times \mathbf{v}\|$ *is the area of the parallelogram.*
4. *From* (1.15)*, observe that for the inner product*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta,$$

which is the projection of one vector onto the other. Similarly, for the cross product

$$\mathbf{u} \times \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \sin \theta \, \mathbf{n},$$

where $\mathbf{n}$ *is a unit vector normal to the plane containing* $\mathbf{u}$ *and* $\mathbf{v}$*. The cross product is a measure of rotation of some sort. For example, it is typically first encountered in physics and mechanics when computing moments or torques produced by forces acting about an axis.*

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

1.6.2 Vectors in Mechanics – Caution

In our mathematics courses, we learn that vectors can be moved freely throughout a coordinate system as long as they maintain their same magnitude and direction. This is why we learned to add vectors by placing the "tail" of one vector at the "tip" of the other as illustrated in Figure 1.9 to give the resultant vector $\mathbf{u} + \mathbf{v}$.

In mechanics, however, where we deal with forces and moments expressed as vectors, we must be more careful. In statics and dynamics of *rigid* – non-deformable – bodies, forces and moments must be applied along their "line of action," which is the infinitely long line on which the vector resides. As illustrated in Figure 1.10, pushing and pulling on a *rigid* body are the same. This is known as the *principle of transmissibility*.

In the mechanics of deformable bodies, such as strength of materials, force and moment vectors must be applied at their "point of action." That is, they can no longer be moved freely along their line of action. In other words, pushing and pulling on a deformable body, such as a string, are not the same as illustrated in Figure 1.11. In summary, mathematically, vectors can be moved freely in three-dimensional space as long as they maintain the same magnitude and direction. In rigid-body mechanics, they must maintain the same magnitude, direction, and *line* of action. In deformable-body mechanics, they must maintain the same magnitude, direction, and *point* of action.

1.6.3 *N*-Dimensional Vectors

The inner product and norm operations, but not the cross product, extend naturally to N-dimensional vectors; for example, the inner product of two N-dimensional vectors

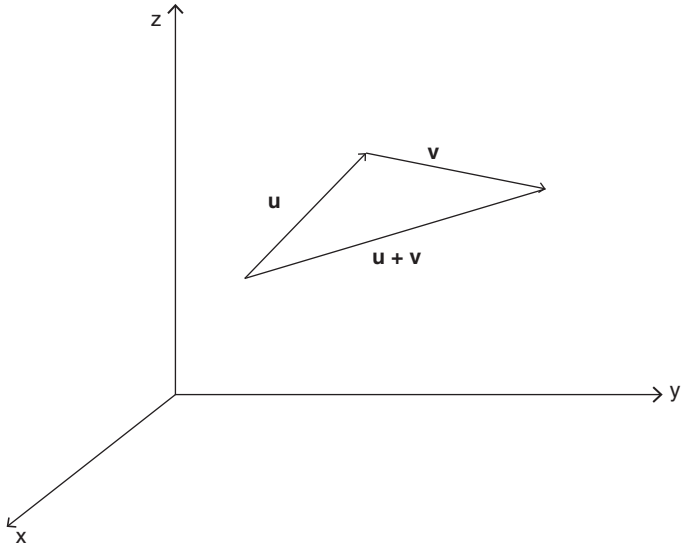


Figure 1.9 Vector addition.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

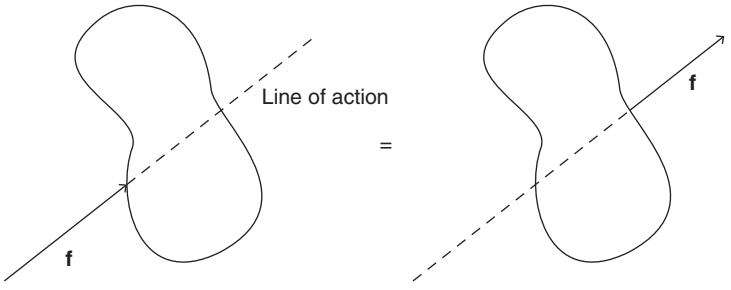


Figure 1.10 Force vector must act along its line of action in rigid-body statics and dynamics.

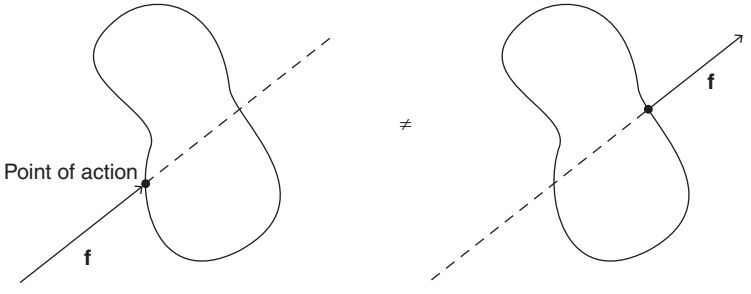


Figure 1.11 Force vector must act at its point of action in deformable-body mechanics of solids.

is given by

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2 + \cdots + u_N v_N,$$

and the norm[13] is

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = \sqrt{u_1^2 + u_2^2 + \cdots + u_N^2};$$

however, graphical interpretations do not.

Given the term "inner" product, one may wonder if there is an "outer" product. Indeed there is; however, we will not find much practical use for it (see Sections 2.8 and 4.2 for two exceptions). Whereas the inner product of $\mathbf{u}$ and $\mathbf{v}$ is $\mathbf{u}^T\mathbf{v}$, which produces a scalar (1×1 matrix), the outer product is $\mathbf{u}\mathbf{v}^T$, which produces an $M \times N$ matrix, where $\mathbf{u}$ is $M \times 1$ and $\mathbf{v}$ is $N \times 1$ ($M = N$ for there to be an inner product). The outer product is also called a *dyad*.

1.7 Vector Spaces, Bases, and Orthogonalization

Mathematically, the primary use of vectors is in providing a *basis* – a generalization of a coordinate system – for a *vector space*. For example, the orthogonal unit vectors

[13] Note that $\mathbf{u}^2$ is taken to mean the inner product of $\mathbf{u}$ with itself, that is, $\langle \mathbf{u}, \mathbf{u} \rangle$; however, one must be careful using this notation as $\sqrt{\mathbf{u}^2} \neq \mathbf{u}$ because $\mathbf{u}^2$ is a scalar.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

i, **j**, and **k** provide a basis for three-dimensional Cartesian coordinates, such that any three-dimensional vector can be written as a linear combination of the basis vectors.

1.7.1 Linear Independence of Vectors

In order to form the *basis* for a vector space, the basis vectors must be *linearly independent*. Essentially, this means that none of the vectors can be written as linear combinations of the others.

In two dimensions, if two vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are such that constants a_1 and a_2 exist for which

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 = \mathbf{0},$$

where a_1 and a_2 are not both zero, then (1) $\mathbf{u}_1$ is a scalar multiple of $\mathbf{u}_2$, (2) $\mathbf{u}_1$ and $\mathbf{u}_2$ are parallel, and (3) $\mathbf{u}_1$ and $\mathbf{u}_2$ are *linearly dependent*. If no such a_1 and a_2 exist, and the preceding expression is only satisfied if $a_1 = a_2 = 0$, then $\mathbf{u}_1$ and $\mathbf{u}_2$ are *linearly independent*.

Similarly, in N dimensions, a set of M vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M$ are *linearly independent* if their linear combination is zero, in which case

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_M\mathbf{u}_M = \mathbf{0} \tag{1.16}$$

only if $a_1 = a_2 = \cdots = a_M = 0$. In other words, none of the vectors may be expressed as a linear combination of the others. If the a_m coefficients are not all zero, then the $\mathbf{u}_m$ vectors are *linearly dependent*.

To obtain a criterion for the existence of the a_m coefficients, take the inner product of each vector $\mathbf{u}_m, m = 1, 2, \ldots, M$ with (1.16), which produces the following system of equations for the coefficients $\mathbf{a}_m, m = 1, 2, \ldots, M$:

$$\begin{aligned}
a_1\mathbf{u}_1^2 + a_2\langle\mathbf{u}_1, \mathbf{u}_2\rangle + \cdots + a_M\langle\mathbf{u}_1, \mathbf{u}_M\rangle &= 0,\\
a_1\langle\mathbf{u}_2, \mathbf{u}_1\rangle + a_2\mathbf{u}_2^2 + \cdots + a_M\langle\mathbf{u}_2, \mathbf{u}_M\rangle &= 0,\\
&\vdots\\
a_1\langle\mathbf{u}_M, \mathbf{u}_1\rangle + a_2\langle\mathbf{u}_M, \mathbf{u}_2\rangle + \cdots + a_M\mathbf{u}_M^2 &= 0,
\end{aligned}$$

or in matrix form

$$\begin{bmatrix}
\mathbf{u}_1^2 & \langle\mathbf{u}_1, \mathbf{u}_2\rangle & \cdots & \langle\mathbf{u}_1, \mathbf{u}_M\rangle\\
\langle\mathbf{u}_2, \mathbf{u}_1\rangle & \mathbf{u}_2^2 & \cdots & \langle\mathbf{u}_2, \mathbf{u}_M\rangle\\
\vdots & \vdots & \ddots & \vdots\\
\langle\mathbf{u}_M, \mathbf{u}_1\rangle & \langle\mathbf{u}_M, \mathbf{u}_2\rangle & \cdots & \mathbf{u}_M^2
\end{bmatrix}
\begin{bmatrix} a_1\\ a_2\\ \vdots\\ a_M \end{bmatrix} =
\begin{bmatrix} 0\\ 0\\ \vdots\\ 0 \end{bmatrix}.$$

The coefficient matrix $\mathbf{G}$ is called the *Gram*, or *Gramian*, *matrix*. If $G = |\mathbf{G}| \neq 0$, where G is the *Gram determinant*, $\mathbf{G}$ is nonsingular and invertible, and the only solution of the homogeneous system is $\mathbf{a} = \mathbf{0}$, which is the trivial solution. In this case, the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_M$ are linearly independent. If $G = |\mathbf{G}| = 0$, an infinity of solutions exist for the coefficients, that is, there is no unique solution, and the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_M$ are linearly dependent.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

REMARKS:

1. *The Gram matrix* $\mathbf{G}$ *is symmetric owing to the properties of inner products.*
2. *If the matrix* $\mathbf{A}$ *is formed by placing the vectors* $\mathbf{u}_m, m = 1, \ldots, M$ *as its columns, then* $\mathbf{G} = \mathbf{A}^T\mathbf{A}$*, which will feature prominently in Sections 2.4 and 2.8.*
3. *Taking the rows (or columns) of a matrix as a set of vectors, the number of linearly independent row (or column) vectors equals the rank of the matrix.*

Example 1.10 Consider the case when $\mathbf{u}_1, \ldots, \mathbf{u}_M$ are all mutually orthogonal, and determine if the vectors are linearly independent.

Solution

Because the vectors are orthogonal, their inner products are all zero except along the main diagonal, in which case

$$\begin{bmatrix} \|\mathbf{u}_1\|^2 & 0 & \cdots & 0 \\ 0 & \|\mathbf{u}_2\|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \|\mathbf{u}_M\|^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus, $G \neq 0$ and the only solution is the trivial solution $\mathbf{a} = \mathbf{0}$, and $\mathbf{u}_1, \ldots, \mathbf{u}_M$ are linearly independent. As one would expect, orthogonal vectors are linearly independent.

1.7.2 Basis of a Vector Space

Consider the linearly independent, N-dimensional vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_R$. All of the vectors $\mathbf{v}$ that are linear combinations of these vectors $\mathbf{u}_m$, such that

$$\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_R\mathbf{u}_R,$$

where the a_m coefficients are arbitrary, form a *vector space* V with *dimension* R. We say that (1) V is a subspace of N-dimensional space; (2) the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_R$ *span* V and form a *basis* for it; (3) if $\mathbf{u}_m$ are unit vectors, each a_m is the *component* of $\mathbf{v}$ along $\mathbf{u}_m$; (4) R is the rank of the Gram matrix $\mathbf{G}$ with defect equal to $N - R$; (5) R is also the rank of the matrix with the $\mathbf{u}_m$ vectors as its rows or columns; and (6) the *range* of a matrix $\mathbf{A}$, denoted by range($\mathbf{A}$), is the vector space spanned by the columns of $\mathbf{A}$.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Example 1.11 Determine whether the vectors

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

span three-dimensional space.

Solution

This is the same as asking if an arbitrary three-dimensional vector $\mathbf{v}$ can be expressed as a linear combination of $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$? That is,

$$\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + a_3\mathbf{u}_3,$$

or

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + a_3 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix},$$

or

$$\begin{aligned} v_1 &= a_1 + a_2 + 2a_3, \\ v_2 &= a_1 + a_3, \\ v_3 &= 2a_1 + a_2 + 3a_3, \end{aligned}$$

or

$$\underbrace{\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}.$$

Note that the columns of $\mathbf{A}$ are $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$.

Is this linear system *consistent* for all $\mathbf{v}$? To find out, evaluate the determinant (only possible if $\mathbf{A}$ is square)

$$|\mathbf{A}| = 0 + 2 + 2 - 0 - 3 - 1 = 0.$$

Because the determinant is zero, the system is inconsistent, and no unique solution exists for $\mathbf{a}$ with any $\mathbf{v}$. Hence, the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ do not span all of three-dimensional space.

Alternatively, we could determine the Gram matrix $\mathbf{G}$ and evaluate the Gram determinant as follows:

$$\mathbf{G} = \begin{bmatrix} \mathbf{u}_1^2 & \langle \mathbf{u}_1, \mathbf{u}_2 \rangle & \langle \mathbf{u}_1, \mathbf{u}_3 \rangle \\ \langle \mathbf{u}_2, \mathbf{u}_1 \rangle & \mathbf{u}_2^2 & \langle \mathbf{u}_2, \mathbf{u}_3 \rangle \\ \langle \mathbf{u}_3, \mathbf{u}_1 \rangle & \langle \mathbf{u}_3, \mathbf{u}_2 \rangle & \mathbf{u}_3^2 \end{bmatrix},$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

which in this case is

$$\mathbf{G} = \begin{bmatrix} 6 & 3 & 9 \\ 3 & 2 & 5 \\ 9 & 5 & 14 \end{bmatrix}.$$

Observe that $\mathbf{G}$ is symmetric as expected. The Gram determinant is $G = |\mathbf{G}| = 0$; therefore, $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ are not linearly independent and cannot span all of three-dimensional space.

Alternatively, we could use Gaussian elimination to row reduce the system to determine if $\text{rank}(\mathbf{A}) = N$. In this example, $\text{rank}(\mathbf{A}) < N$ because the third equation is the sum of the first and second equations. Observe that it is easiest to do the first method (if $\mathbf{A}$ is square), but that either of the other two methods is required for nonsquare $\mathbf{A}$.

In contrast to the preceding example, if the three three-dimensional basis vectors are linearly independent, they span all of three-dimensional vector space. As such, any three-dimensional vector can be written as a linear combination of these three basis vectors. This holds in general as well. That is, an N-dimensional vector space is spanned by *any* set of N linearly independent, N-dimensional vectors.

Summarizing: If $\mathbf{A}$ is an $N \times N$ matrix, the following statements are equivalent:

1. $|\mathbf{A}| \neq 0$.
2. $\mathbf{A}$ is *nonsingular*.
3. $\mathbf{A}$ is *invertible*.
4. $\mathbf{A}$ has *rank* $\text{rank}(\mathbf{A}) = N$.
5. The row (and column) vectors of $\mathbf{A}$ are *linearly independent*.
6. The row (and column) vectors of $\mathbf{A}$ *span* N-dimensional space and form a *basis* for it.
7. $\mathbf{Au} = \mathbf{0}$ has only the *trivial* solution, which is $\mathbf{u} = \mathbf{A}^{-1}\mathbf{0} = \mathbf{0}$.
8. $\mathbf{Au} = \mathbf{b}$ is *consistent* for every $\mathbf{b}$, with $\mathbf{u} = \mathbf{A}^{-1}\mathbf{b}$, and there is a unique solution for each $\mathbf{b}$.

Note that the opposite of (1) through (7), but not (8), is also equivalent, in which case the system of equations would either be *inconsistent* – having no solution – or the system would have an infinity of solutions.

1.7.3 Gram–Schmidt Orthogonalization

Although any set of N linearly independent, N-dimensional vectors can form a basis for an N-dimensional vector space, it is often convenient to work with basis vectors that are mutually orthogonal and of unit length. We say that they are *orthonormal* as they are orthogonal vectors normalized to length one. This can be accomplished using *Gram–Schmidt orthogonalization*, which produces an orthogonal set of unit vectors $(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_S)$ from S linearly independent vectors $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_S)$. Then

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

any N-dimensional vector $\mathbf{v}$ in the vector space V spanned by $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_S$ can be written as a linear combination of the orthonormal basis vectors as follows:

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{q}_1 \rangle \, \mathbf{q}_1 + \langle \mathbf{v}, \mathbf{q}_2 \rangle \, \mathbf{q}_2 + \cdots + \langle \mathbf{v}, \mathbf{q}_S \rangle \, \mathbf{q}_S,$$

where the inner product $\langle \mathbf{v}, \mathbf{q}_i \rangle$ is the component of $\mathbf{v}$ in the $\mathbf{q}_i$ direction.

Procedure (using $\mathbf{u}_1$ as the reference vector):

1. Normalize $\mathbf{u}_1$ according to

$$\mathbf{q}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|},$$

   such that $\mathbf{q}_1$ is a unit vector.
2. Determine the component of $\mathbf{u}_2$ that is orthogonal to $\mathbf{q}_1$. This is accomplished by subtracting the component of $\mathbf{u}_2$ in the $\mathbf{q}_1$ direction from $\mathbf{u}_2$ to obtain

$$\hat{\mathbf{u}}_2 = \mathbf{u}_2 - \langle \mathbf{u}_2, \mathbf{q}_1 \rangle \, \mathbf{q}_1,$$

   which is orthogonal to $\mathbf{q}_1$ as illustrated in Figure 1.12. Normalizing $\hat{\mathbf{u}}_2$ gives

$$\mathbf{q}_2 = \frac{\hat{\mathbf{u}}_2}{\|\hat{\mathbf{u}}_2\|}.$$

3. Determine the component of $\mathbf{u}_3$ that is orthogonal to both $\mathbf{q}_1$ *and* $\mathbf{q}_2$. This is accomplished by subtracting the components of $\mathbf{u}_3$ in the $\mathbf{q}_1$ and $\mathbf{q}_2$ directions from $\mathbf{u}_3$ as follows:

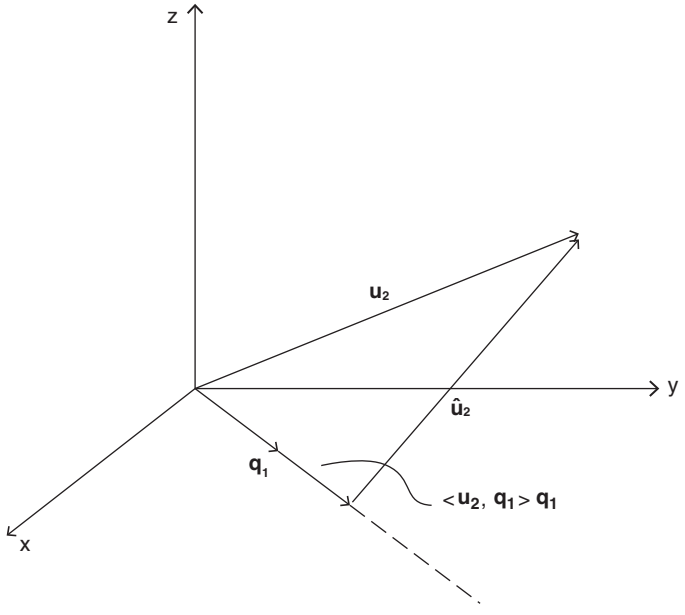


Figure 1.12 Step 2 in Gram–Schmidt orthogonalization.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

$$\hat{\mathbf{u}}_3 = \mathbf{u}_3 - \langle \mathbf{u}_3, \mathbf{q}_1 \rangle \, \mathbf{q}_1 - \langle \mathbf{u}_3, \mathbf{q}_2 \rangle \, \mathbf{q}_2,$$

thereby producing a vector that is orthogonal to both $\mathbf{q}_1$ and $\mathbf{q}_2$. Normalizing $\hat{\mathbf{u}}_3$ produces

$$\mathbf{q}_3 = \frac{\hat{\mathbf{u}}_3}{\|\hat{\mathbf{u}}_3\|}.$$

4. Continue for $4 \leq i \leq S$:

$$\hat{\mathbf{u}}_i = \mathbf{u}_i - \sum_{k=1}^{i-1} \langle \mathbf{u}_i, \mathbf{q}_k \rangle \, \mathbf{q}_k, \quad \mathbf{q}_i = \frac{\hat{\mathbf{u}}_i}{\|\hat{\mathbf{u}}_i\|}. \tag{1.17}$$

Let us confirm that Gram–Schmidt orthogonalization in the form of (1.17) does indeed produce a vector $\hat{\mathbf{u}}_i$ that is orthogonal to each of the previously obtained orthonormal vectors $\mathbf{q}_j, j = 1, \ldots, i-1$. To do so, we show that the inner products of $\hat{\mathbf{u}}_i$ and $\mathbf{q}_j$ are all zero. This inner product is

$$\langle \hat{\mathbf{u}}_i, \mathbf{q}_j \rangle = \langle \mathbf{u}_i, \mathbf{q}_j \rangle - \sum_{k=1}^{i-1} \langle \mathbf{u}_i, \mathbf{q}_k \rangle \langle \mathbf{q}_k, \mathbf{q}_j \rangle, \quad j = 1, \ldots, i-1.$$

But $\langle \mathbf{q}_k, \mathbf{q}_j \rangle = 0$ for $j \neq k$, and $\langle \mathbf{q}_k, \mathbf{q}_j \rangle = 1$ when $j = k$, thereby producing

$$\langle \hat{\mathbf{u}}_i, \mathbf{q}_j \rangle = \langle \mathbf{u}_i, \mathbf{q}_j \rangle - \langle \mathbf{u}_i, \mathbf{q}_j \rangle = 0.$$

Consequently, $\hat{\mathbf{u}}_i$ is orthogonal to each of the previously obtained vectors $\mathbf{q}_j, j = 1, \ldots, i-1$.

REMARKS:

1. *The Gram–Schmidt orthogonalization procedure always results in orthonormal basis vectors if the* $\mathbf{u}_i$ *vectors are linearly independent, but changing the order of the linearly independent vectors* $\mathbf{u}_i$ *will alter the resulting orthonormal vectors* $\mathbf{q}_i$*. In other words, Gram–Schmidt orthogonalization is not unique.*
2. *It will be found that this orthogonalization procedure bears a close relationship to QR decomposition (see Section 2.10) and least-squares methods (see Section 10.2).*

1.7.4 Row, Column, and Null Spaces of a Matrix

Because the rows and columns of a matrix form a set of vectors, it is possible, and often useful, to consider the vector space spanned by the row or column vectors of a matrix, that is, the vector space for which the row or column vectors form a basis. The *row space* of matrix $\mathbf{A}$ is the vector space spanned by its row vectors, and the *column space*, or *range*, is that spanned by its column vectors.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Let us consider the $M \times N$ matrix $\mathbf{A}$ of the form

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{bmatrix},$$

where $\mathbf{u}_m, m = 1, \ldots, M$, are the row vectors of $\mathbf{A}$, and $\mathbf{v}_n, n = 1, \ldots, N$, are the column vectors of $\mathbf{A}$. The row representation of $\mathbf{A}$ is then

$$\mathbf{A} = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_M^T \end{bmatrix}, \quad \text{where} \quad \mathbf{u}_m = \begin{bmatrix} A_{m1} \\ A_{m2} \\ \vdots \\ A_{mN} \end{bmatrix}, m = 1, \ldots, M. \tag{1.18}$$

Similarly, the column representation of $\mathbf{A}$ is

$$\mathbf{A} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_N \end{bmatrix}, \tag{1.19}$$

where

$$\mathbf{v}_n = \begin{bmatrix} A_{1n} \\ A_{2n} \\ \vdots \\ A_{Mn} \end{bmatrix}, \quad n = 1, \ldots, N.$$

Let us consider the system of linear algebraic equations

$$\mathbf{A}\mathbf{u} = \mathbf{b}. \tag{1.20}$$

Suppose $\mathbf{u}_H$ is a solution of the homogeneous system

$$\mathbf{A}\mathbf{u}_H = \mathbf{0}, \tag{1.21}$$

and $\mathbf{u}_P$ is a particular solution of the nonhomogeneous system

$$\mathbf{A}\mathbf{u}_P = \mathbf{b}. \tag{1.22}$$

Because this is a linear system of algebraic equations, the general solution is of the form

$$\mathbf{u} = \mathbf{u}_P + k\mathbf{u}_H,$$

where k is an arbitrary constant. In order to show this, observe that

$$\mathbf{A}(\mathbf{u}_P + k\mathbf{u}_H) = \mathbf{b}$$
$$\mathbf{A}\mathbf{u}_P + k\mathbf{A}\mathbf{u}_H = \mathbf{b}$$
$$\mathbf{b} + k\mathbf{0} = \mathbf{b},$$

which is of course true for all $\mathbf{b}$. We say that the homogeneous solution $\mathbf{u}_H$ is in the *null space*, and $\mathbf{b}$ is in the *column space*, of the matrix $\mathbf{A}$ as further clarified later in this section.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Let us first consider the homogeneous equation (1.21) using the row representation as follows:

$$\begin{bmatrix} \cdots & \mathbf{u}_1^T & \cdots \\ \cdots & \mathbf{u}_2^T & \cdots \\ & \vdots & \\ \cdots & \mathbf{u}_M^T & \cdots \end{bmatrix} \begin{bmatrix} \vdots \\ \mathbf{u}_H \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \mathbf{0} \\ \vdots \end{bmatrix}.$$

This requires that

$$\langle \mathbf{u}_1, \mathbf{u}_H \rangle = 0, \quad \langle \mathbf{u}_2, \mathbf{u}_H \rangle = 0, \quad \ldots, \quad \langle \mathbf{u}_M, \mathbf{u}_H \rangle = 0;$$

that is, $\mathbf{u}_H$ is orthogonal to each of the row vectors of $\mathbf{A}$. The following possibilities then exist:

1. If the rank of $\mathbf{A}$ is N [rank($\mathbf{A}$) $= N$], then the row vectors $\mathbf{u}_m, m = 1, \ldots, M$, are linearly independent, and they span the N-dimensional row space. In this case, $\mathbf{u}_H = \mathbf{0}$ is the only solution of $\mathbf{A}\mathbf{u}_H = \mathbf{0}$, and the null space is the zero vector.
2. If rank($\mathbf{A}$) $< N$, in which case only rank($\mathbf{A}$) of the $\mathbf{u}_m$, $m = 1, \ldots, M$ vectors are linearly independent, then the rank($\mathbf{A}$)-dimensional space spanned by the rank($\mathbf{A}$) linearly independent row vectors is the *row space*. There are then [N − rank($\mathbf{A}$)] arbitrary vectors that satisfy $\mathbf{A}\mathbf{u}_H = \mathbf{0}$ and are thus orthogonal to the row space. The [N − rank($\mathbf{A}$)]-dimensional space where these $\mathbf{x}_H$ vectors exist is called the *orthogonal complement* of the row space, or the *null space*, of $\mathbf{A}$. It follows that a nonzero vector $\mathbf{u}_H$ satisfies $\mathbf{A}\mathbf{u}_H = \mathbf{0}$ *if and only if* it is in the null space of matrix $\mathbf{A}$.

Now consider the nonhomogeneous equation (1.20) using the column representation as follows:

$$\begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_N \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} = \begin{bmatrix} \vdots \\ \mathbf{b} \\ \vdots \end{bmatrix}.$$

This requires that vector $\mathbf{b}$ be a linear combination of the column vectors $\mathbf{v}_n, n = 1, \ldots, N$, in which case

$$u_1\mathbf{v}_1 + u_2\mathbf{v}_2 + \cdots + u_N\mathbf{v}_N = \mathbf{b}.$$

Then $\mathbf{b}$ is in the column space of $\mathbf{A}$, and the column space defines all possible vectors $\mathbf{b}$. The following possibilities then exist:

1. If $\mathbf{b}$ is in the column space of $\mathbf{A}$, then the system $\mathbf{A}\mathbf{u} = \mathbf{b}$ is consistent and has a unique solution if rank($\mathbf{A}$) $= N$ or an infinity of solutions if rank($\mathbf{A}$) $< N$.
2. If $\mathbf{b}$ is not in the column space of $\mathbf{A}$, then the system $\mathbf{A}\mathbf{u} = \mathbf{b}$ is inconsistent.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

REMARKS:

1. *The dimensions of the row and column space are the same as one another and equal the rank of matrix* $\mathbf{A}$*, that is,* $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T)$.
2. *For an* $M \times N$ *matrix, the dimension of the null space is equal to* $N - \text{rank}(\mathbf{A})$.
3. *If one considers the matrix* $\mathbf{A}$ *to be a linear transformation from an* N*-dimensional vector space to an* M*-dimensional one, then the column space of the matrix equals the image of this linear transformation.*
4. *Elementary row operations do not change the row or column space of a matrix.*

Our first exposure to linear algebra is typically in the context of solving systems of linear algebraic equations for which a unique solution is sought. In such cases, the coefficient matrix is typically square and the rank equals the number of unknowns. Indeed, only square matrices have a determinant, can be invertible and have eigenvalues and eigenvectors (see Chapter 2). As a result, it is tempting to dismiss discussion of rectangular matrices, or more precisely those having $\text{rank}(\mathbf{A}) \neq N$, as only being of interest to mathematicians. Being comfortable with vector spaces and the particulars of matrices that are not square will pay dividends when we get to applications, such as least-squares methods, singular-value decomposition, and proper-orthogonal decomposition, for example. Rest assured that we would not include these topics if not relevant to applications of interest to scientists and engineers.

1.8 Linear Transformations

Thus far, we have viewed the matrix problem $\mathbf{Au} = \mathbf{b}$ as a system of linear algebraic equations, in which we seek the solution vector $\mathbf{u}$ for a given coefficient matrix $\mathbf{A}$ and right-hand-side vector $\mathbf{b}$. Alternatively, the matrix problem

$$\underset{M\times N}{\mathbf{A}}\ \underset{N\times 1}{\mathbf{x}} = \underset{M\times 1}{\mathbf{y}},$$

where $\mathbf{A}$ is specified and $\mathbf{x}$ and $\mathbf{y}$ are arbitrary vectors, may be viewed as a *linear transformation*, or *mapping*, from an N-dimensional vector space to an M-dimensional vector space, which are the sizes of $\mathbf{x}$ and $\mathbf{y}$, respectively. We say that the matrix $\mathbf{A}$ operates on the vector $\mathbf{x}$ to transform it into the vector $\mathbf{y}$. In this context, $\mathbf{A}$ transforms vectors from an N-space *domain* to an M-space *range* as shown in Figure 1.13. Such transformation matrices will be encountered in numerous settings throughout the remainder of the text, and thinking of matrices as linear operators in this manner will prove helpful in understanding many of the methods to come.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

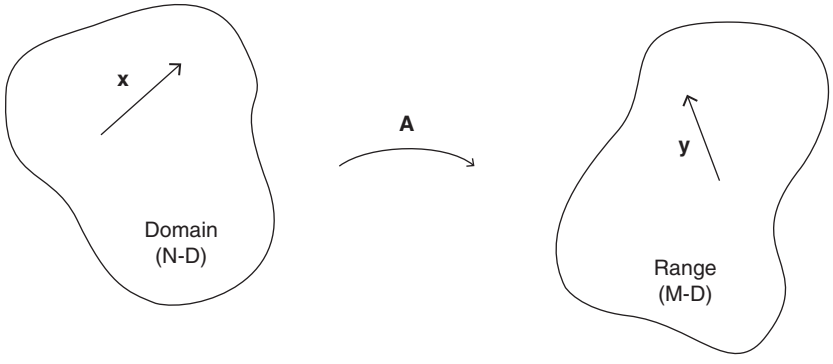


Figure 1.13 Linear transformation from **x** in the domain to **y** in the range of **A**.

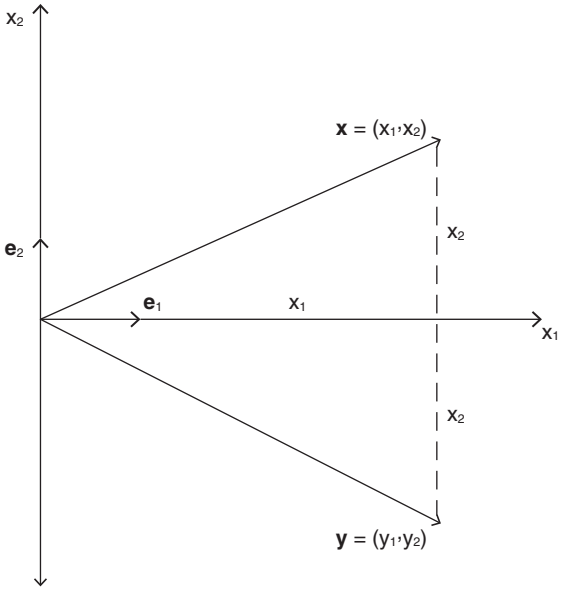


Figure 1.14 Reflection of vector **x** about the x_1-axis to form vector **y** in Example 1.12

Example 1.12 Determine the transformation matrix **A** that transforms the vector **x** to the vector **y** by reflecting it about the x_1-axis as illustrated in Figure 1.14.

Solution

The vector **y** is given by

$$\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2,$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are unit vectors in the x_1 and x_2 coordinate directions, respectively. As can be observed graphically, however, the components of $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T$ are related to the components of $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ according to

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

$$y_1 = x_1, \quad y_2 = -x_2,$$

for such a reflection. These two relationships can be expressed as a matrix transformation of the form

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Example 1.13 Determine the transformation matrix $\mathbf{A}$ that transforms the vector $\mathbf{x}$ to the vector $\mathbf{y}$ by rotating it counterclockwise about the origin through an angle θ as illustrated in Figure 1.15. Both vectors are of length r.

Solution

The vector $\mathbf{x}$ is given by

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = r\cos\alpha\ \mathbf{e}_1 + r\sin\alpha\ \mathbf{e}_2,$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are unit vectors in the x_1 and x_2 coordinate directions, respectively. Thus,

$$x_1 = r\cos\alpha, \quad x_2 = r\sin\alpha. \tag{1.23}$$

Similarly, the vector $\mathbf{y}$ is given by

$$\begin{aligned} \mathbf{y} &= y_1\mathbf{e}_1 + y_2\mathbf{e}_2, \\ &= r\cos(\alpha+\theta)\mathbf{e}_1 + r\sin(\alpha+\theta)\mathbf{e}_2, \\ \mathbf{y} &= r\left(\cos\alpha\cos\theta - \sin\alpha\sin\theta\right)\mathbf{e}_1 + r\left(\sin\alpha\cos\theta + \cos\alpha\sin\theta\right)\mathbf{e}_2. \end{aligned}$$

From (1.23), however, this can be written in terms of x_1 and x_2 as follows:

$$\mathbf{y} = \left(x_1\cos\theta - x_2\sin\theta\right)\mathbf{e}_1 + \left(x_2\cos\theta + x_1\sin\theta\right)\mathbf{e}_2.$$

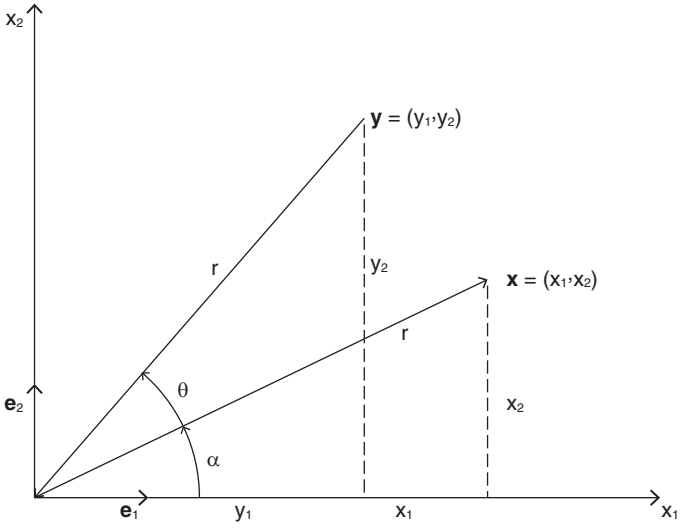


Figure 1.15 Rotation of vector $\mathbf{x}$ through an angle θ to form vector $\mathbf{y}$ in Example 1.13.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Therefore,

$$y_1 = x_1 \cos\theta - x_2 \sin\theta, \quad y_2 = x_1 \sin\theta + x_2 \cos\theta,$$

which in matrix form is

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

This transformation matrix rotates a vector counterclockwise about the origin through an angle θ without altering its length.

REMARKS:

1. *Because the transformations are linear, they can be superimposed. For example,*

$$\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x} = \mathbf{y}$$

is equivalent to

$$(\mathbf{A} + \mathbf{B})\,\mathbf{x} = \mathbf{y}.$$

2. *Successively applying a series of transformations to* $\mathbf{x}$*, such as*

$$\mathbf{x}_1 = \mathbf{A}_1\mathbf{x}, \quad \mathbf{x}_2 = \mathbf{A}_2\mathbf{x}_1, \quad \ldots, \quad \mathbf{x}_K = \mathbf{A}_K\mathbf{x}_{K-1},$$

is equivalent to applying the transformation

$$\mathbf{x}_K = \mathbf{A}\mathbf{x},$$

where $\mathbf{A} = \mathbf{A}_K \cdots \mathbf{A}_2\mathbf{A}_1$ *(note the reverse order).*

3. *If the transformation matrix* $\mathbf{A}$ *is invertible, then* $\mathbf{A}^{-1}\mathbf{y}$ *transforms* $\mathbf{y}$ *back to* $\mathbf{x}$*, which is the inverse transformation.*
4. *If* $\mathbf{y} = \mathbf{A}\mathbf{x}$*, then* $\mathbf{y}$ *is a linear combination of the columns* $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N)$ *of* $\mathbf{A}$ *in the form*

$$\mathbf{y} = \mathbf{A}\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_N\mathbf{a}_N.$$

That is, $\mathbf{y}$ *is in the column space of matrix* $\mathbf{A}$*.*

5. *In the case when a matrix* $\mathbf{A}$ *transforms a nonzero vector* $\mathbf{x}$ *to the zero vector* $\mathbf{0}$*, that is,* $\mathbf{A}\mathbf{x} = \mathbf{0}$*, the vector* $\mathbf{x}$ *is said to be in the null space of* $\mathbf{A}$*. The null space is a subspace of that defined by the row or column vectors of* $\mathbf{A}$ *(see Section 1.7.4).*
6. *If the transformation matrix* $\mathbf{A}$ *is orthogonal, it preserves the length of the vector, that is,* $\|\mathbf{x}\| = \|\mathbf{y}\|$*. It is shown in Section 2.4 that the rotation matrix in Example 1.13 is orthogonal.*
7. *For more on the mathematics of linear algebra, see Strang (2006), Strang (2016), Lay, Lay, and McDonald (2016), and Meckes and Meckes (2018).*

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

1.9 Note on Norms

Throughout this chapter, we have made use of the norm as a measure of the size or length of a vector. It essentially replaces the absolute value of real numbers or the modulus of complex numbers. Thus far, the norm has been defined for an N-dimensional vector $\mathbf{u}$ as

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = \sqrt{u_1^2 + u_2^2 + \cdots + u_N^2} = \left(\sum_{n=1}^{N} u_n^2 \right)^{1/2},$$

which is the square root of the inner product of the vector with itself. In two and three dimensions, this definition of the norm corresponds to the Euclidian length of the vector – the straight-line distance from the origin to the point (u_1, u_2, u_3). It turns out that while this is the most common definition of the norm, there are others that are encountered as well.

In general, we define the so-called L_p-norm, or simply the p-norm, of the vector $\mathbf{u}$ as follows:

$$\|\mathbf{u}\|_p = \left(\sum_{n=1}^{N} |u_n|^p \right)^{1/p}, \tag{1.24}$$

where $1 \leq p < \infty$. Therefore, our previous definition of the norm corresponds to the L_2-norm, or simply the "2-norm." As a result, many authors indicate this norm by $\|\mathbf{u}\|_2$. We will use this notation when necessary to prevent any confusion; however, if no subscript is given, then it is assumed to be the L_2-norm.

Other common norms correspond to $p = 1$ and $p = \infty$. The L_1-norm is the sum of the absolute values of each component of $\mathbf{u}$ according to

$$\|\mathbf{u}\|_1 = \sum_{n=1}^{N} |u_n|.$$

This corresponds to only being able to travel from the origin to the point along lines parallel to the coordinate axes. The L_∞-norm, or ∞-norm, is given by

$$\|\mathbf{u}\|_\infty = \max_{1 \leq n \leq N} |u_n|,$$

which is simply the largest component of $\mathbf{u}$ by absolute value.

We will also have occasion to compute norms of matrices in order to quantify their "size." The L_1-norm of a matrix $\mathbf{A}$ is the largest L_1-norm of the column vectors of the $M \times N$ matrix $\mathbf{A}$ as follows:

$$\|\mathbf{A}\|_1 = \max_{1 \leq n \leq N} \|\mathbf{a}_n\|_1,$$

where $\mathbf{a}_n$ denotes the nth column of $\mathbf{A}$. The L_∞-norm of a matrix $\mathbf{A}$ is the largest L_1-norm of the row vectors of $\mathbf{A}$, that is,

$$\|\mathbf{A}\|_\infty = \max_{1 \leq m \leq M} \|\mathbf{a}_m\|_1,$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

where $\mathbf{a}_m$ denotes the mth row of $\mathbf{A}$. Therefore, the 1-norm is the maximum column sum, and the ∞-norm is the maximum row sum of the matrix $\mathbf{A}$. Finally, the L_2-norm of a matrix is given by

$$\|\mathbf{A}\|_2 = \left[\rho\left(\mathbf{A}^T\mathbf{A}\right)\right]^{1/2} = \sigma_1,$$

where $\rho\left(\mathbf{A}^T\mathbf{A}\right)$ is the spectral radius, that is, the largest eigenvalue by magnitude, of $\mathbf{A}^T\mathbf{A}$ (see Section 2.2), and σ_1 is the maximum singular value of $\mathbf{A}$ (see Section 2.8). Hence, it is also sometimes referred to as the *spectral norm*. If $\mathbf{A}$ is symmetric, then the L_2-norm is the spectral radius of $\mathbf{A}$, that is, $\|\mathbf{A}\|_2 = \rho(\mathbf{A})$.

To more directly mimic the L_2-norm for vectors, we also define the *Frobenius norm* of a matrix given by

$$\|\mathbf{A}\|_F = \left(\sum_{m=1}^{M}\sum_{n=1}^{N}|A_{mn}|^2\right)^{1/2}.$$

This is simply the square root of the sum of the squares of all the elements of the matrix $\mathbf{A}$. Alternatively, it can also be expressed in terms of $\mathbf{A}\mathbf{A}^T$ as

$$\|\mathbf{A}\|_F = \sqrt{\text{tr}(\mathbf{A}\mathbf{A}^T)} = \left(\sum_{m=1}^{M}\lambda_m\right)^{1/2} = \left(\sum_{m=1}^{M}\sigma_m^2\right)^{1/2},$$

where $\text{tr}(\mathbf{A}\mathbf{A}^T)$ is the trace of $\mathbf{A}\mathbf{A}^T$, λ_m are the eigenvalues of $\mathbf{A}\mathbf{A}^T$, and σ_m are the singular values of $\mathbf{A}$ (see Chapter 2).

REMARKS:

1. *When both vectors and matrices are present, we use the same norm for both when performing operations.*
2. *In one dimension, the L_2-norm is simply the absolute value.*
3. *Unless indicated otherwise, we will use the L_2-norm for both vectors and matrices. Its popularity stems in part from the fact that it can often be interpreted in terms of energies in certain applications. In addition, it extends naturally to functions as in Chapter 3.*
4. *Although the L_2-norm is most commonly used, the L_1-norm and L_∞-norm are typically more convenient to compute.*
5. *The Schwarz inequality for vectors is*

   $$|\langle\mathbf{u},\mathbf{v}\rangle| \leq \|\mathbf{u}\|_2\|\mathbf{v}\|_2.$$

   More generally, it can be proven that

   $$\|\mathbf{A}\mathbf{B}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|.$$

   These inequalities prove useful in determining bounds on various quantities of interest involving vectors and matrices.
6. *Vector and matrix norms will play a prominent role in defining the condition number in Chapter 6 that is so central to computational methods.*

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

7. *For more on the mathematical properties of norms, see Golub and Van Loan (2013) and Horn and Johnson (2013).*

1.10 Briefly on Bases

Let us end this chapter where we started, with the choice of coordinate systems and vector bases. Early in our training as scientists and engineers, we learn that the key to solving many problems is a judicious choice of coordinate system, or vector basis, with respect to which to represent the problem. Do we employ a Cartesian, cylindrical, or spherical coordinate system? Where is the origin of the coordinate system logically placed? How are the coordinate directions oriented? These choices are somewhat arbitrary and are simply a matter of convenience in many cases, while in other situations it is the difference between being able to obtain a solution or not. We will see over and over again that this important lesson holds true in more advanced pursuits as well.

The reason why this focus on bases is so helpful is that more than half the battle of solving a problem is typically knowing what form, or basis, the solution can most naturally be written with respect to. In other words, if we know what vectors or functions the solution can be written in terms of, then we just need to determine how much of each basis is required for the particular solution at hand. The simplest example of this is the basis vectors of a coordinate system. For example, we know that any vector in three-dimensional Cartesian coordinates can be written as a linear combination of the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ in the three coordinate directions. For a given vector, it then remains to simply determine how much of each unit vector is required to represent that particular vector – that is, to determine the components u_x, u_y, and u_z in the vector

$$\mathbf{u} = u_x\mathbf{i} + u_y\mathbf{j} + u_z\mathbf{k}.$$

We espoused the virtues of using orthonormal basis vectors to provide a convenient representation of a problem and ease many of the operations, such as projecting vectors onto coordinate axes and transforming vectors. The Gram–Schmidt orthogonalization procedure was introduced as a means to obtain such a basis from a set of linearly independent basis vectors.

It will also prove helpful to view solutions of various types of problems in terms of linear combinations of basis vectors or functions. For example, we know that the solution of a linear, constant-coefficient ordinary differential equation is given in terms of the functions e^{mx}, where there are as many m's as the order of the differential equation. The solution of such a differential equation then involves simply determining the m's and how much of each e^{mx} term is required. For a second-order linear ordinary differential equation with constant coefficients, therefore, we seek the constants m_1, m_2, c_1, c_2 in the function

$$u(x) = c_1e^{m_1x} + c_2e^{m_2x}.$$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

In all such scenarios, then, the two-step process involves answering the following questions:

1. How do we choose or determine the set of basis vectors or functions to use for a given problem?
2. Given the set of basis vectors or functions, how do we determine the coefficients for each?

As we develop a variety of techniques for solving and analyzing the various types of matrix and differential problems and their solutions, it is helpful to take note of the following questions as they relate to bases:

1. For the given form of coefficient or transformation matrix $\mathbf{A}$, what can we say about the form of the solution vectors or eigenvectors? Can we represent the solution in terms of a particular set of basis vectors? For example, can we use orthonormal basis vectors or must we employ a more general set of linearly independent, but not orthogonal, basis vectors?
2. Can we modify the basis vectors in order to facilitate obtaining the solution or ease interpretation?

Such questions related to bases will be a common theme throughout the text as they greatly help in understanding the motivation as well as relative advantages and disadvantages of various solution techniques and numerical methods. This being the case, the reader is strongly encouraged to become comfortable with the terminology and techniques surrounding basis vectors (and eventually functions). To aid in this process, we will close each of the first four chapters in Part I with a "Briefly on Bases" section that highlights what that chapter has contributed to the preceding discussion. In so doing, the reader will have a solid foundation on which to build in Part III, where these ideas are central to methods for analyzing and understanding large-scale numerical solutions and experimental data sets.

Finally, the notational conventions used throughout the text are summarized in Table 1.2.

Exercises

The following exercises are to be completed using hand calculations. When appropriate, the results can be checked using built-in functions in Python, MATLAB, or Mathematica for the vector and matrix operations.

1.1 Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

Determine $\mathbf{A}^T$, $\mathbf{A}^2$, $|\mathbf{A}|$, and $\mathbf{A}^{-1}$.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

Table 1.2 Notational conventions.

Quantity/operation	Convention	Examples
Vectors	Bold lowercase letters and square brackets	$\mathbf{a} = [\cdot]$
Matrices	Bold uppercase letters and square brackets	$\mathbf{A} = [\cdot]$
Zero vector or matrix	Bold zero	$\mathbf{0}$
Identity matrix	Bold I	$\mathbf{I}$
Matrix/array size	Uppercase letters	I, J, K, M, N
Matrix/array index	Lowercase letters and subscripts	i, j, k, m, n
Trace of matrix		$\text{tr}(\mathbf{A})$
Determinant		$\lvert\mathbf{A}\rvert$
Condition number		$\text{cond}(\mathbf{A})$, $\kappa(\mathbf{A})$
Spectral radius		$\rho(\mathbf{A})$
Matrix rank		$\text{rank}\,(\mathbf{A})$
Vector or matrix norm		$\lVert\mathbf{a}\rVert$, $\lVert\mathbf{A}\rVert$
Matrix transpose	Superscript T	$\mathbf{A}^T$
Matrix inverse	Superscript -1	$\mathbf{A}^{-1}$
Matrix pseudo-inverse	Superscript $+$	$\mathbf{A}^{+}$
Scalar/dot/inner product	Angle brackets	$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b}$
Outer product		$\mathbf{a}\mathbf{b}^T$
Inner product of functions	Angle brackets	$\langle g, h \rangle = \int_a^b g(x)h(x)dx$
Functions	Lowercase with independent variables in parentheses	$f(x,t)$
Functionals	Uppercase with dependent functions in square brackets	$I[u(x), v(x)], J[u(t)]$
Spatial ordinary derivative	Prime, superscript in parentheses	$u'(x) = du/dx$, $u^{(n)}(x) = d^n u/dx^n$
Temporal ordinary derivative	Dot	$\dot{u}(t) = du/dt, \ddot{u}(t) = d^2u/dt^2$
Partial derivatives	Subscripts	$u_{xx} = \partial^2 u/\partial x^2$
Eigenvalues	Lowercase lambda	λ
Lagrange multipliers	Uppercase lambda	Λ
Gradient operator	Del (nabla)	$\boldsymbol{\nabla}$
Laplacian operator	Del (nabla) squared	$\boldsymbol{\nabla}^2 = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$
General differential operator		$\mathcal{L}$
Adjoint differential operator		$\mathcal{L}^*$
Imaginary number	San serif i	i
Real and imaginary parts		$\text{Re}(u)$, $\text{Im}(u)$
Complex conjugate	Overbar	$\overline{\mathbf{A}}$
Numerical time step		Δt
Numerical grid spacing		$\Delta x, \Delta y$
Numerical grid point	Index subscript(s)	$u_{i,j}$
Numerical iteration number	Superscript in parentheses	$u^{(n)}$

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

1.2 Consider the two vectors and one matrix

$$\mathbf{u} = \begin{bmatrix} -3 \\ 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ 1 \\ 3 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 5 & 4 & -2 \\ -7 & 2 & 1 & 4 \\ 3 & 2 & -1 & 1 \\ 1 & 3 & -4 & 2 \end{bmatrix}.$$

Determine $\mathbf{u} + \mathbf{v}$, $\langle \mathbf{u}, \mathbf{v} \rangle$, $\mathbf{A}^T$, $|\mathbf{A}|$, $\mathbf{A}^{-1}$, and $\mathbf{Au}$.

1.3 Show that

$$|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$$

if $\mathbf{A}$ and $\mathbf{B}$ are general 2×2 matrices.

1.4 For a general 3×3 matrix $\mathbf{A}$, write out the cofactor expansion of the determinant of $\mathbf{A}$ in the form

$$|\mathbf{A}| = A_{11}C_{11} + A_{12}C_{12} + A_{13}C_{13},$$

and reorder the terms to show that

$$|\mathbf{A}| = \left|\mathbf{A}^T\right|.$$

1.5 The *law of cosines* for a triangle with sides of length a, b, and c, in which the angle opposite the side of length c is C, takes the form

$$c^2 = a^2 + b^2 - 2ab\cos C.$$

Prove this by taking vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ along each side, such that $\mathbf{c} = \mathbf{a} - \mathbf{b}$, and considering the inner product $\langle \mathbf{c}, \mathbf{c} \rangle = \langle \mathbf{a} - \mathbf{b}, \mathbf{a} - \mathbf{b} \rangle$.

1.6 Let $\mathbf{x}$ and $\mathbf{y}$ be N-dimensional vectors and λ be a scalar. Prove that

$$\|\mathbf{x} + \lambda\mathbf{y}\|^2 + \|\mathbf{x} - \lambda\mathbf{y}\|^2 = 2\left(\|\mathbf{x}\|^2 + \lambda^2\|\mathbf{y}\|^2\right).$$

1.7 Let $\mathbf{x}$ and $\mathbf{y}$ be N-dimensional orthogonal vectors. Prove that the *Pythagoras theorem* takes the form

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

1.8 Consider the determinant

$$P(\lambda) = \begin{vmatrix} 4-\lambda & 0 & 1 \\ 1 & -\lambda & 1 \\ -1 & -2 & 2-\lambda \end{vmatrix},$$

where λ is an unknown parameter. By expanding the determinant, show that $P(\lambda)$ is a polynomial for λ. Determine the value(s) of λ for which $P(\lambda) = 0$.

1.9 Determine the values of λ for which the following system of equations has a non-trivial solution:

$$\begin{aligned} 3u_1 + u_2 - \lambda u_3 &= 0, \\ 4u_1 - 2u_2 - 3u_3 &= 0, \\ 2\lambda u_1 + 4u_2 + \lambda u_3 &= 0. \end{aligned}$$

For each possible value of λ, determine the most general solution for $\mathbf{u}$.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

1.10 Determine the rank of the matrix

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ -2 & 1 & 1 \end{bmatrix}.$$

1.11 For the matrices

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 1 \\ 1 & 4 & 0 \\ 2 & 1 & -3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & -3 & 1 \\ 2 & 0 & 5 \\ 3 & 1 & 2 \end{bmatrix},$$

verify that

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

1.12 Prove that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ if $\mathbf{A}$ and $\mathbf{B}$ are invertible and of the same size.

1.13 Consider the following system of equations expressed in matrix form:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}.$$

Determine whether this system has no solution, one unique solution, or an infinity of solutions. Justify your answer.

1.14 Consider the following system of equations:

$$\begin{aligned} u_1 + 3u_2 + 2u_3 &= 0, \\ u_1 + u_2 + u_3 &= 1, \\ 2u_2 + u_3 &= 2. \end{aligned}$$

Determine whether this system has no solution, one unique solution, or an infinity of solutions. Justify your answer.

1.15 Given the system of equations

$$\begin{aligned} u_1 + 5u_2 &= 3, \\ 7u_1 - 3u_2 &= -1, \end{aligned}$$

determine the solution using Cramer's rule.

1.16 Given the system of equations

$$\begin{aligned} 2u_1 + u_3 &= 9, \\ 2u_2 + u_3 &= 3, \\ u_1 + 2u_3 &= 3, \end{aligned}$$

determine the solution using Cramer's rule.

1.17 Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 1 & 0 \\ 2 & 1 & 3 & 1 \\ 5 & 6 & 7 & 2 \\ 2 & 1 & 0 & 1 \end{bmatrix}.$$

Determine the solution of the homogeneous linear system of equations $\mathbf{Au} = \mathbf{0}$.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

1.18 Consider the system of linear algebraic equations $\mathbf{Au} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \\ 0 & 3 & 5 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 0 \\ 4 \\ 1 \end{bmatrix}.$$

Determine whether the system is consistent or inconsistent. If it is consistent, obtain its solution(s) for $\mathbf{u}$.

1.19 Consider the system of linear algebraic equations

$$\begin{aligned} 2u_1 - u_2 - u_3 &= 2, \\ u_1 + 2u_2 + u_3 &= 2, \\ 4u_1 - 7u_2 - 5u_3 &= 2. \end{aligned}$$

Show that this system of equations has a one-parameter family of solutions by considering the ranks of relevant matrices. Determine this solution.

1.20 For the system of linear algebraic equations $\mathbf{Au} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 2 & -4 & 1 \\ 6 & 2 & -1 \\ -2 & 6 & -2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 10 \\ -6 \end{bmatrix},$$

obtain the solution $\mathbf{u}$ using (1) Gaussian elimination, (2) matrix inverse, and (3) Cramer's rule.

1.21 In Example 1.1, Ohm's and Kirchhoff's laws were applied to a parallel electric circuit involving voltage sources and resistors to obtain the system of linear algebraic equations $\mathbf{Au} = \mathbf{b}$ of the form

$$\begin{bmatrix} 30 & -10 & -8 \\ -10 & 16 & -6 \\ -8 & -6 & 18 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 6 \end{bmatrix}.$$

Solve the system of equations for the three currents using (1) Gaussian elimination, (2) matrix inverse, and (3) Cramer's rule.

1.22 Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Is the matrix $\mathbf{A}$ singular or nonsingular? Does the system of equations $\mathbf{Au} = \mathbf{0}$ have a nontrivial solution? Are the column vectors of $\mathbf{A}$ linearly dependent?

1.23 Are the vectors $\mathbf{u}_1 = \begin{bmatrix} 2 & 0 & 1 \end{bmatrix}^T$ and $\mathbf{u}_2 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T$ linearly independent? Are they orthogonal? Explain your answers.

1.24 Determine whether the vectors $\mathbf{u}_1 = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}^T$, $\mathbf{u}_2 = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^T$, and $\mathbf{u}_3 = \begin{bmatrix} 1 & 3 & 2 \end{bmatrix}^T$ span three-dimensional space.

1.25 Determine whether the vectors $\mathbf{u}_1 = \begin{bmatrix} -1 & 0 & -2 \end{bmatrix}^T$, $\mathbf{u}_2 = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}^T$ and $\mathbf{u}_3 = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^T$ are linearly independent.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press

1.26 Determine if the vector $\mathbf{v} = \begin{bmatrix} 6 & 1 & -6 & 2 \end{bmatrix}^T$ is in the vector space spanned by the basis vectors $\mathbf{u}_1 = \begin{bmatrix} 1 & 1 & -1 & 1 \end{bmatrix}^T$, $\mathbf{u}_2 = \begin{bmatrix} -1 & 0 & 1 & 1 \end{bmatrix}^T$, and $\mathbf{u}_3 = \begin{bmatrix} 1 & -1 & -1 & 0 \end{bmatrix}^T$.

1.27 Consider the following vectors in three-dimensional Cartesian coordinates:

$$\mathbf{u}_1 = -\mathbf{i} + \mathbf{k},$$
$$\mathbf{u}_2 = 2\mathbf{j} + \mathbf{k},$$
$$\mathbf{u}_3 = \mathbf{i} + \mathbf{j} + \mathbf{k}.$$

Use Gram–Schmidt orthogonalization to obtain a set of orthonormal basis vectors for three-dimensional vector space.

1.28 Using Gram–Schmidt orthogonalization, construct a set of three mutually orthogonal unit vectors from the set of linearly independent vectors $\mathbf{u}_1 = \begin{bmatrix} -1 & 2 & 0 \end{bmatrix}^T$, $\mathbf{u}_2 = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}^T$, and $\mathbf{u}_3 = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}^T$.

1.29 Using Gram–Schmidt orthogonalization, construct a set of three mutually orthogonal unit vectors from the set of linearly independent vectors $\mathbf{u}_1 = \begin{bmatrix} 1 & 0 & 2 & 2 \end{bmatrix}^T$, $\mathbf{u}_2 = \begin{bmatrix} 1 & 1 & 0 & 1 \end{bmatrix}^T$, and $\mathbf{u}_3 = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}^T$.

1.30 Show that the vector $\mathbf{v} = \begin{bmatrix} 2 & 1 & 2 & 0 \end{bmatrix}^T$ is in the vector space spanned by the three vectors given in Exercise 1.29, and express $\mathbf{v}$ as a linear combination of the orthonormal vectors $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$ that result from Gram–Schmidt orthogonalization.

https://doi.org/10.1017/9781108782333.003 Published online by Cambridge University Press