Chapter 8

Systems of Linear Algebraic Equations; Gauss Elimination

8.1 Introduction

There are many applications in science and engineering where application of the relevant physical law(s) immediately produces a set of linear algebraic equations. For instance, the application of Kirchoff's laws to a DC electrical circuit containing any number of resistors, batteries, and current loops immediately produces such a set of equations on the unknown currents. In other cases, the problem is stated in some other form such as one or more ordinary or partial differential equations, but the solution method eventually leads us to a system of linear algebraic equations. For instance, to find a particular solution to the differential equation

$$y''' - y'' = 3x^2 + 5\sin x \tag{1}$$

by the method of undetermined coefficients (Section 3.7.2), we seek it in the form

$$y_p(x) = Ax^4 + Bx^3 + Cx^2 + D\sin x + E\cos x. \tag{2}$$

Putting (2) into (1) and equating coefficients of like terms on both sides of the equation gives five linear algebraic equations on the unknown coefficients $A, B, \ldots, E$. Or, solving the so-called Laplace partial differential equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{3}$$

on the rectangle $0 < x < 1, 0 < y < 1$ by the method of finite differences (which is studied in Section 20.5), using a mesh size $\Delta x = \Delta y = 0.05$, gives $19^2 = 361$ linear algebraic equations on the unknown values of u at the 361 nodal points of the mesh. Our point here is not to get ahead of ourselves by plunging into partial differential equations, but to say that the solution of practical problems of interest in science and engineering *often* leads us to systems of linear algebraic equations.

Such systems often involve a great many unknowns. Thus, the question of existence (Does a solution exist?), which often sounds "too theoretical" to the practicing engineer, takes on great practical importance because a considerable computational effort is at stake.

The subject of linear algebra and matrices encompasses a great deal more than the theory and solution of systems of linear algebraic equations, but the latter is indeed a central topic and is foundational to others. Thus, we begin this sequence of five chapters (8–12) on **linear algebra** with an introduction to the theory of systems of linear algebraic equations, and their solution by the method of Gauss elimination. Results obtained here are used, and built upon, in Chapters 9–12.

Chapters 9 and 10 take us from vectors in 3-space to vectors in n-space and generalized vector space, to matrices and determinants. Linear systems of algebraic equations are considered again, in the second half of Chapter 10, in terms of rank, inverse matrix, LU decomposition, Cramer's rule, and linear transformation. Chapter 11 introduces the eigenvalue problem, diagonalization, and quadratic forms; areas of application include systems of ordinary differential equations, vibration theory, chemical kinetics, and buckling. Chapter 12 is optional and brief and provides an extension of results in Chapters 9–11 to complex spaces.

8.2 Preliminary Ideas and Geometrical Approach

The problem of finding solutions of equations of the form

$$f(x) = 0 \tag{1}$$

occupies a place of both practical and historical importance. Equation (1) is said to be **algebraic**, or **polynomial**, if $f(x)$ is expressible in the form $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where $a_n \neq 0$ for definiteness [i.e., if $f(x)$ is a polynomial of finite degree n], and it is said to be **transcendental** otherwise.

EXAMPLE 1. The equations $6x - 5 = 0$ and $3x^4 - x^3 + 2x + 1 = 0$ are algebraic, whereas $x^3 + 2\sin x = 0$ and $e^x - 3 = 0$ are transcendental since $\sin x$ and e^x cannot be expressed as polynomials of finite degree. ▮

Besides the algebraic versus transcendental distinction, we classify (1) as **linear** if $f(x)$ is a first-degree polynomial,

$$a_1 x + a_0 = 0, \tag{2}$$

and **nonlinear** otherwise. Thus, the first equation in Example 1 is linear, and the other three are nonlinear.

While (1) is one equation in one unknown, we often encounter problems involving more than one equation and/or more than one unknown – that is, a **system** of equations consisting of m equations in n unknowns, where $m \geq 1$ and $n \geq 1$,

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0, \\ f_2(x_1, \ldots, x_n) &= 0, \\ &\vdots \\ f_m(x_1, \ldots, x_n) &= 0 \end{aligned} \tag{3}$$

such as

$$\begin{aligned} x_1 - \sin(x_1 + 7x_2) &= 0, \\ x_2^3 + x_2 - 5x_1 + 6 &= 0. \end{aligned} \tag{4}$$

In (4) it happens that $m = n$ (namely, $m = n = 2$) so that there are as many equations as unknowns. In general, however, m may be less than, equal to, or greater than n so we allow for $m \neq n$ in this discussion even though $m = n$ is the most important case.

In this chapter we consider only the case where (3) is **linear**, of the form

$$\boxed{\begin{array}{llll} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = c_1, & \text{(eq.1)} \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = c_2, & \text{(eq.2)} \\ \qquad\qquad\vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = c_m, & \text{(eq.m)} \end{array}} \tag{5}$$

and we restrict m and n to be finite, and the a_{ij}'s and c_j's to be real numbers. If all the c_j's are zero then (5) is **homogeneous**; if they are not all zero then (5) is **nonhomogeneous**.

The subscript notation adopted in (5) is not essential but is helpful in holding the nomenclature to a minimum, in rendering inherent patterns more visible, and in permitting a natural transition to matrix notation. The first subscript in a_{ij} indicates the equation, and the second indicates the x_j variable that it multiplies. For instance, a_{21} appears in the second equation and multiplies the x_1 variable. To avoid ambiguity we should write $a_{2,1}$ rather than a_{21} so that one does not mistakenly read the subscripts as twenty-one, but we will omit commas except when such ambiguity is not easily resolved from the context.

We say that a sequence of numbers $s_1, s_2, \ldots, s_n$ is a **solution** of (5) if and only if each of the m equations is satisfied numerically when we substitute s_1 for x_1, s_2 for x_2, and so on. If there exist *one or more* solutions to (5), we say that the system is **consistent**; if there is precisely *one* solution, that solution is **unique**; and if there is more than one, the solution is **nonunique**. If, on the other hand, there are *no* solutions to (5), the system is said to be **inconsistent**. The collection of all solutions to (5) is called its **solution set** so, by "solving (5)" we mean finding its solution set.

Let us begin with the simple case, where $m = n = 1$:

$$a_{11}x_1 = c_1. \tag{6}$$

In the generic case, $a_{11} \neq 0$ and (6) admits the unique solution $x_1 = c_1/a_{11}$, but if $a_{11} = 0$ there are two possibilities: if $c_1 \neq 0$ then there are no values of x_1 such that $0x_1 = c_1$ and (6) is inconsistent, but if $c_1 = 0$ then (6) becomes $0x_1 = 0$, and $x_1 = \alpha$ is a solution for *any* value of α; that is, the solution is nonunique.

Far from being too simple to be of interest, the case where $m = n = 1$ establishes a pattern that will hold in general, for any values of m and n. Specifically, the system (5) will admit a unique solution, no solution, or an infinity of solutions. For instance, it will never admit 4 solutions, 12 solutions, or 137 solutions.

Next, consider the case where $m = n = 2$:

$$a_{11}x_1 + a_{12}x_2 = c_1, \qquad \text{(eq.1)} \tag{7a}$$

$$a_{21}x_1 + a_{22}x_2 = c_2. \qquad \text{(eq.2)} \tag{7b}$$

If a_{11} and a_{12} are not both zero, then (eq.1) defines a straight line, say $L1$, in a Cartesian x_1, x_2 plane; that is the solution set of (eq.1) is the set of all points on that line. Similarly, if a_{21} and a_{22} are not both zero then the solution set of (eq.2) is the set of all points on a straight line $L2$. There exist exactly three possibilities, and these are illustrated in Fig. 1. First, the lines may intersect at a point, say P, in which case (7) admits the unique solution given by the coordinate pair x_1, x_2 of the point P (Fig. 1a). That is, any solution pair x_1, x_2 of (7) needs to be in the solution set of (eq.1) *and* in the solution set of (eq.2) hence at an intersection of $L1$ and $L2$. This is the generic case, and it occurs (Exercise 2) as long as

$$a_{11}a_{22} - a_{12}a_{21} \neq 0; \tag{8}$$

(8) is the analog of the $a_{11} \neq 0$ condition for the $m = n = 1$ case discussed above.

Second, the lines may be parallel and nonintersecting (Fig. 1b), in which case there is no solution. Then (7) is inconsistent, the solution set is empty.

Third, the lines may coincide (Fig. 1c), in which case the coordinate pair of each point on the line is a solution. Then (7) is consistent and there are an infinite number of solutions.

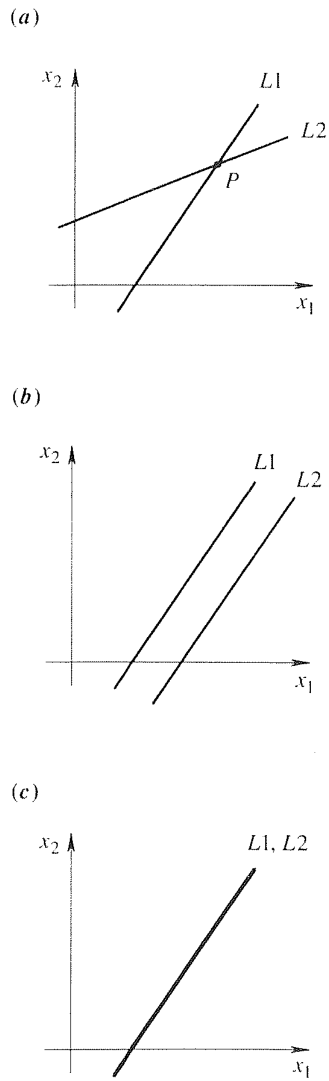


Figure 1. Existence and uniqueness for the system (7).

EXAMPLE 2.

$$\begin{aligned} 2x_1 - x_2 &= 5, \\ x_1 + 3x_2 &= -1, \end{aligned} \qquad \begin{aligned} x_1 + 3x_2 &= 1, \\ x_1 + 3x_2 &= 0, \end{aligned} \qquad \begin{aligned} x_1 + 3x_2 &= 1, \\ 2x_1 + 6x_2 &= 2, \end{aligned}$$

illustrate these three cases, respectively. ▮

Below (7) we said "If a_{11} and a_{12} are not both zero . . . ." What if they are both zero? Then if $c_1 \neq 0$ there is no solution of (7a), and hence there is no solution to the system (7). But if $c_1 = 0$, then (7a) reduces to $0 = 0$ and can be discarded, leaving just (7b). If a_{21} and a_{22} are not both zero, then (7b) gives a line of solutions, but if they are both zero then everything hinges on c_2. If $c_2 \neq 0$ there is no solution and (7) is inconsistent, but if $c_2 = 0$, so both (7a) and (7b) are simply $0 = 0$, then both x_1 and x_2 are arbitrary, and every point in the plane is a solution.

Next, consider the case where $m = n = 3$:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = c_1, \qquad \text{(eq.1)} \qquad \text{(9a)}$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = c_2, \qquad \text{(eq.2)} \qquad \text{(9b)}$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = c_3. \qquad \text{(eq.3)} \qquad \text{(9c)}$$

Continuing the geometric approach exemplified by Fig. 1, observe that if a_{11}, a_{12}, a_{13} are not all zero then (eq.1) defines a plane, say $P1$, in Cartesian x_1, x_2, x_3 space, and similarly for (eq.2) and (eq.3). In the generic case, $P1$ and $P2$ intersect along a line L, and L pierces $P3$ at a point P. Then the x_1, x_2, x_3 coordinates of P give the unique solution of (9).

In the nongeneric case we can have no solution or an infinity of solutions in the following ways. There will be no solution if L is parallel to $P3$ and hence fails to pierce it, or if any two of the planes are parallel and not coincident. There will be an infinity of solutions if L lies in $P3$ (i.e., a line of solutions), if two planes are coincident and intersect the third (again, a line of solutions), or if all three planes are coincident (this time an entire plane of solutions).

The case where all of the a_{ij} coefficients are zero in one or more of equations (9) is left for the exercises.

An abstract extension of such geometrical reasoning can be continued even if $m = n \geq 4$. For instance, one speaks of $a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = c_1$ as defining a *hyperplane* in an abstract four-dimensional space. In fact, perhaps we should mention that even the familiar x_1, x_2 plane and x_1, x_2, x_3 space discussed here could be abstract as well. For instance, if x_1 and x_2 are unknown currents in two loops of an electrical circuit, then what physical meaning is there to an x_1, x_2 plane? None, but we can introduce it, create it, to assist our reasoning.

Closure. Most of this section is devoted to a geometrical discussion of the system (5) of linear algebraic equations. A great advantage of geometrical reasoning is that it brings our visual system into play. It is estimated that at least a third of the neurons in our brain are devoted to vision, hence our visual sense is extremely sophisticated. No wonder we say "Now I see what you mean; now I get the picture." The more geometry, pictures, visual images to aid our thinking, the better! We have not yet aimed at theorems, and have been content to lay the groundwork for the ideas of existence and uniqueness of solutions. In considering the cases where $m = n = 1$, $m = n = 2$, and $m = n = 3$, we have not meant to imply that we need to have $m = n$; *all* possibilities are considered in the next section. To proceed further, we need to consider the process of *finding* solutions, and that we do, in Section 8.3, by the method of Gauss elimination.

EXERCISES 8.2

1. True or false? If false, give a counterexample.

(a) An algebraic equation is necessarily linear.
(b) An algebraic equation is necessarily nonlinear.
(c) A transcendental equation is necessarily linear.
(d) A transcendental equation is necessarily nonlinear.
(e) A linear equation is necessarily algebraic.
(f) A nonlinear equation is necessarily algebraic.
(g) A linear equation is necessarily transcendental.
(h) A nonlinear equation is necessarily transcendental.

2. Derive the condition (8) as the necessary and sufficient condition for (7) to admit a unique solution.

3. (a) Discuss all possibilities of the existence and uniqueness of solutions of (9) from a geometrical point of view, in the event that $a_{11} = a_{12} = a_{13} = 0$, but a_{21}, a_{22}, a_{23} and a_{31}, a_{32}, a_{33} are not all zero.
(b) Same as (a), but with $a_{21} = a_{22} = a_{23} = 0$ as well.
(c) Same as (a), but with $a_{21} = a_{22} = a_{23} = a_{31} = a_{32} = a_{33} = 0$ as well.

8.3 Solution by Gauss Elimination

8.3.1. Motivation. In this section we continue to consider the system of m linear algebraic equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= c_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= c_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= c_m, \end{aligned} \tag{1}$$

in the n unknowns $x_1, \ldots, x_n$, and develop the solution technique known as **Gauss elimination**. To motivate the ideas, we begin with an example.

EXAMPLE 1. Determine the solution set of the system

$$\begin{aligned} x_1 + x_2 - x_3 &= 1, \\ 3x_1 + x_2 + x_3 &= 9, \\ x_1 - x_2 + 4x_3 &= 8. \end{aligned} \tag{2}$$

Keep the first equation intact, and add -3 times the first equation to the second (as a replacement for the second equation), and add -1 times the first equation to the third (as a replacement for the third equation). These steps yield the new "indented" system

$$\begin{aligned} x_1 + x_2 - x_3 &= 1, \\ -2x_2 + 4x_3 &= 6, \\ -2x_2 + 5x_3 &= 7. \end{aligned} \tag{3}$$

Next, keep the first two of these intact, and add -1 times the second equation to the third, and obtain

$$\begin{aligned} x_1 + x_2 - x_3 &= 1, \\ -2x_2 + 4x_3 &= 6, \\ x_3 &= 1. \end{aligned} \tag{4}$$

Finally, multiplying the second of these by $-1/2$ to normalize the leading coefficient (to unity), gives

$$\begin{aligned} x_1 + x_2 - x_3 &= 1, && \text{(eq.1)} \\ x_2 - 2x_3 &= -3, && \text{(eq.2)} \\ x_3 &= 1. && \text{(eq.3)} \end{aligned} \tag{5}$$

It is helpful to think of the original system (2) as a tangle of string that we wish to unravel. The first step is to find a loose end and that is, in effect, what the foregoing process of successive indentations has done for us. Specifically, (eq.3) in (5) is the "loose end," and with that in hand we may unravel (5) just as we would unravel a tangle: putting $x_3 = 1$ into (eq.2) gives $x_2 = -1$, and then putting $x_3 = 1$ and $x_2 = -1$ into (eq.1) gives $x_1 = 3$. Thus, we obtain the unique solution

$$x_3 = 1, \quad x_2 = -1, \quad x_1 = 3. \tag{6}$$

COMMENT 1. From a mathematical point of view, the system (2) was a "tangle" because the equations were **coupled**; that is, each equation contained more than one unknown. Actually, the final system (5) is coupled too, since (eq.1) contains all three unknowns and (eq.2) contains two of them. However, the coupling in (5) is not as debilitating because (5) is in what we call **triangular form**. Thus, we were able to solve (eq.3) for x_3, put that value into (eq.2) and solve for x_2, and then put these values into (eq.1) and solve for x_1, which steps are known as **back substitution**.

COMMENT 2. However, the process begs this question: Is it obvious that the systems (2)–(5) all have the same solution sets so that when we solve (5) we are actually solving (2)? That is, is it not conceivable that in applying the arithmetic steps that carried us from (2) to (5) we might, inadvertently, have altered the solution set? For example, $x - 1 = 4$ has the unique solution $x = 5$, but if we innocently square both sides, the resulting equation $(x-1)^2 = 16$ admits the *two* solutions $x = 5$ and $x = -3$. ▮

The question just raised applies to linear systems in general. It is answered in Theorem 8.3.1 that follows, but first we define two terms: "equivalent systems" and "elementary equation operations."

Two linear systems in n unknowns, x_1 through x_n, are said to be **equivalent** if their solution sets are identical.

The following operations on linear systems are known as **elementary equation operations**:

1. Addition of a multiple of one equation to another
   Symbolically: (eq.j) $\to$ (eq.j) $+ \alpha$ (eq.k)

2. Multiplication of an equation by a nonzero constant
   Symbolically: (eq.j) $\to \alpha$ (eq.j)

3. Interchange of two equations
   Symbolically: (eq.j) $\leftrightarrow$ (eq.k)

Then we can state the following result.

THEOREM 8.3.1 *Equivalent Systems*
If one linear system is obtained from another by a finite number of elementary equation operations, then the two systems are equivalent.

Outline of Proof: The truth of this claim for elementary equation operations of types 2 and 3 should be evident, so we confine our remarks to operations of type 1. It suffices to look at the effect of one such operation. Thus, suppose that a given linear system A is altered by replacing its jth equation by its jth plus α times its kth, its other equations being kept intact. Let us call the new system A'. Surely, every solution of A will also be a solution A' since we have merely added equal quantities to equal quantities. That is, *if A' results from A by the application of an elementary equation operation of type 1, then every solution of A is also a solution of A'*. Further, we can convert A' back to A by an elementary equation operation of type 1, namely, by replacing the jth equation of A' by the jth equation of A' plus $-\alpha$ times the kth equation of A'. Consequently, it follows from the italicized result (two sentences back) that every solution of A' is also a solution of A. Then A and A' are equivalent, as claimed. ■

In Example 1, we saw that each step is an elementary equation operation: Three elementary equation operations of type 1 took us from (2) to (4), and one of type 2 took us from (4) to (5); finally, the back substitution amounted to several operations of type 1. Thus, according to Theorem 8.3.1, equivalence was maintained throughout so we can be sure that (6) is the solution set of the original system (2) (as can be verified by direct substitution).

The system in Example 1 admitted a unique solution. To see how the method of successive elimination works out when there is no solution, or a nonunique solution, let us work two more examples.

EXAMPLE 2. *Inconsistent System.* Consider the system

$$\begin{aligned} 2x_1 + 3x_2 - 2x_3 &= 4, \\ x_1 - 2x_2 + x_3 &= 3, \\ 7x_1 \qquad\quad - x_3 &= 2. \end{aligned} \tag{7}$$

Keep the first equation intact, add $-\frac{1}{2}$ times the first equation to the second (eq.2 $\rightarrow$ eq.2 $-\frac{1}{2}$ eq.1), and add $-\frac{7}{2}$ times the first to the third (eq.3 $\rightarrow$ eq.3 $-\frac{7}{2}$ eq.1):

$$\begin{aligned} 2x_1 + 3x_2 - 2x_3 &= 4, \\ -\tfrac{7}{2}x_2 + 2x_3 &= 1, \\ -\tfrac{21}{2}x_2 + 6x_3 &= -12. \end{aligned} \tag{8}$$

Keep the first two equations intact, and add -3 times the second equation to the third (eq.3

$\rightarrow$ eq.3 -3 eq.2):

$$\begin{aligned} 2x_1 + 3x_2 - 2x_3 &= 4, \\ -\tfrac{7}{2}x_2 + 2x_3 &= 1, \\ 0 &= -15. \end{aligned} \tag{9}$$

Any solution x_1, x_2, x_3 of (9) must satisfy each of the three equations, but there are no values of x_1, x_2, x_3 that can satisfy $0 = -15$. Thus, (9) is inconsistent (has no solution), and therefore (7) is as well.

COMMENT. The source of the inconsistency is the fact that whereas the left-hand side of the third equation is 2 times the left-hand side of the first equation plus 3 times the left-hand side of the second, the right-hand sides do not bear that relationship: $2(4)+3(3) = 17 \neq 2$. [While that built-in contradiction is not obvious from (7), it eventually comes to light in the third equation in (9).] If we modify the system (7) by changing the final 2 in (7) to 17, then the final -12 in (8) becomes a 3, and the final -15 in (9) becomes a zero:

$$\begin{aligned} 2x_1 + 3x_2 - 2x_3 &= 4, \\ -\tfrac{7}{2}x_2 + 2x_3 &= 1, \\ 0 &= 0 \end{aligned} \tag{10}$$

or, multiplying the first by $\frac{1}{2}$ and the second by $-\frac{2}{7}$,

$$\begin{aligned} x_1 + \tfrac{3}{2}x_2 - x_3 &= 2, \\ x_2 - \tfrac{4}{7}x_3 &= -\tfrac{2}{7}, \end{aligned} \tag{11a,b}$$

where we have discarded the identity $0 = 0$. Thus, by changing the c_j's so as to be "compatible," the system now admits an infinity of solutions rather than none. Specifically, we can let x_3 (or x_2, it doesn't matter which) in (11b) be *any* value, say α, where α is arbitrary. Then (11b) gives $x_2 = -\frac{2}{7} + \frac{4}{7}\alpha$, and putting these into (11a), $x_1 = \frac{17}{7} + \frac{1}{7}\alpha$. Thus, we have the infinity of solutions

$$x_3 = \alpha, \quad x_2 = -\frac{2}{7} + \frac{4}{7}\alpha, \quad x_1 = \frac{17}{7} + \frac{1}{7}\alpha \tag{12}$$

for any α. Evidently, two of the three planes intersect, giving a line that lies in the third plane, and equations (12) are parametric equations of that line! ∎

EXAMPLE 3. *Nonunique Solution.* Consider the system of four equations in six unknowns ($m = 4, n = 6$)

$$\begin{aligned} 2x_2 + x_3 + 4x_4 + 3x_5 + x_6 &= 2, \\ x_1 - x_2 + x_3 + 2x_6 &= 0, \\ x_1 + x_2 + 2x_3 + 4x_4 + x_5 + 2x_6 &= 3, \\ x_1 - 3x_2 - 4x_4 - 2x_5 + x_6 &= 0. \end{aligned} \tag{13}$$

Wanting the top equation to begin with x_1 and subsequent equations to indent at the left, let us first move the top equation to the bottom (eq.1 $\leftrightarrow$ eq.4):

$$\begin{aligned} x_1 - 3x_2 \qquad\quad - 4x_4 - 2x_5 + x_6 &= 0, \\ x_1 - x_2 + x_3 \qquad\qquad\quad + 2x_6 &= 0, \\ x_1 + x_2 + 2x_3 + 4x_4 + x_5 + 2x_6 &= 3, \\ 2x_2 + x_3 + 4x_4 + 3x_5 + x_6 &= 2. \end{aligned} \tag{14}$$

Add -1 times the first equation to the second (eq.2 $\to$ eq.2 -1 eq.1) and third (eq.3 $\to$ eq.3 -1 eq.1) equations:

$$\begin{aligned} x_1 - 3x_2 \qquad\quad - 4x_4 - 2x_5 + x_6 &= 0, \\ 2x_2 + x_3 + 4x_4 + 2x_5 + x_6 &= 0, \\ 4x_2 + 2x_3 + 8x_4 + 3x_5 + x_6 &= 3, \\ 2x_2 + x_3 + 4x_4 + 3x_5 + x_6 &= 2. \end{aligned} \tag{15}$$

Add -2 times the second to the third (eq.3 $\to$ eq.3 -2 eq.2) and -1 times the second to the fourth (eq.4 $\to$ eq.4 -1 eq.2):

$$\begin{aligned} x_1 - 3x_2 \qquad\quad - 4x_4 - 2x_5 + x_6 &= 0, \\ 2x_2 + x_3 + 4x_4 + 2x_5 + x_6 &= 0, \\ -x_5 - x_6 &= 3, \\ x_5 \qquad &= 2. \end{aligned} \tag{16}$$

Add the third to the fourth (eq.4 $\to$ eq.4 + eq.3):

$$\begin{aligned} x_1 - 3x_2 \qquad\quad - 4x_4 - 2x_5 + x_6 &= 0, \\ 2x_2 + x_3 + 4x_4 + 2x_5 + x_6 &= 0, \\ -x_5 - x_6 &= 3, \\ -x_6 &= 5. \end{aligned} \tag{17}$$

Finally, multiply the second, third, and fourth by $\frac{1}{2}$, -1, and -1, respectively, to normalize the leading coefficients (eq.2 $\to \frac{1}{2}$ eq.2, eq.3 $\to -1$ eq.3, eq.4 $\to -1$ eq.4):

$$\begin{aligned} x_1 - 3x_2 \qquad\quad - 4x_4 - 2x_5 + x_6 &= 0, \\ x_2 + \tfrac{1}{2}x_3 + 2x_4 + x_5 + \tfrac{1}{2}x_6 &= 0, \\ x_5 + x_6 &= -3, \\ x_6 &= -5. \end{aligned} \tag{18}$$

The last two equations give $x_6 = -5$ and $x_5 = 2$, and these values can be substituted back into the second equation. In that equation we can let x_4 be arbitrary, say α_1, and we can also let x_3 be arbitrary, say α_2. Then that equation gives x_2 and, again by back substitution, the first equation gives x_1. The result is the infinity of solutions

$$\begin{aligned} &x_6 = -5, \quad x_5 = 2, \quad x_4 = \alpha_1, \quad x_3 = \alpha_2, \\ &x_2 = \frac{1}{2} - 2\alpha_1 - \frac{1}{2}\alpha_2, \quad x_1 = \frac{21}{2} - 2\alpha_1 - \frac{3}{2}\alpha_2, \end{aligned} \tag{19}$$

where α_1 and α_2 are arbitrary. ∎

If a solution set contains p independent arbitrary parameters $(\alpha_1, \dots, \alpha_p)$, we call it (in this text) a ***p*-parameter family of solutions**. Thus, (12) and (19) are one- and two-parameter families of solutions, respectively. Each choice of values for $\alpha_1, \dots, \alpha_p$ yields a **particular solution**. In (19), for instance, the choice $\alpha_1 = 1$ and $\alpha_2 = 0$ yields the particular solution $x_1 = \frac{17}{2}$, $x_2 = -\frac{3}{2}$, $x_3 = 0$, $x_4 = 1$, $x_5 = 2$, and $x_6 = -5$.

8.3.2. Gauss elimination. The method of Gauss elimination,* illustrated in Examples 1–3, can be applied to *any* linear system (1), whether or not the system is consistent, and whether or not the solution is unique. Though hard to tell from the foregoing hard calculations, the method is efficient and is commonly available in computer systems.

Observe that the end result of the Gauss elimination process enables us to determine, merely from the pattern of the final equations, whether or not a solution exists and is unique. For instance, we can see from the pattern of (5) that there is a unique solution, from the bottom equation in (9) that there no solution, and from the extra double indentation in (18) that there is a two-parameter family of solutions.

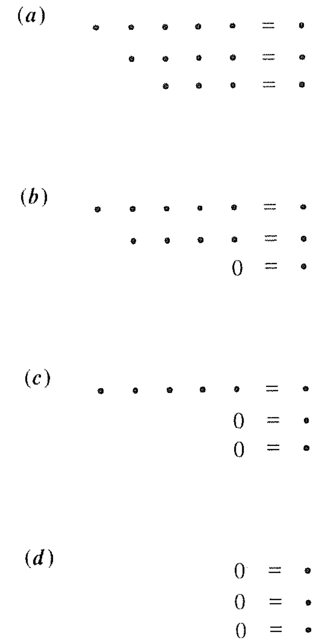


Figure 1. The final pattern; $m = 3, n = 5$.

As representative of the case where $m < n$, let $m = 3$ and $n = 5$. There are four possible final patterns, and these are shown schematically in Fig. 1. For instance, the third equation in Fig. 1a could be $x_3 - 6x_4 + 2x_5 = 0$ or $x_3 + 2x_4 + 0x_5 = 4$, and the given third equation in Fig. 1b could be $0 = 6$ or $0 = 0$. It may seem foolish to include the case shown in Fig. 1d because there are no x_j's (all of the a_{ij} coefficients being zero), but it is *possible* so we have included it. From these patterns we draw these conclusions: (a) there exists a two-parameter family of solutions; (b) there is no solution (the system is inconsistent) if the right-hand member of the third equation is nonzero, and a three-parameter family of solutions if the latter is zero; (c) there is no solution if either of the right-hand members of the second and third equations is nonzero, and a four-parameter family of solutions if each of them is zero; (d) there is no solution if any of the right-hand members is nonzero, and a five-parameter family of solutions if each of them is zero.

It may appear that Fig. 1 does not cover all possible cases. For instance, what about the case shown in Fig. 2? That case can be converted to the case shown in Fig. 1a simply by renaming the unknowns: let x_3 become x_2 and let x_5 become x_3. Specifically, let $x_1 \to x_1$, $x_3 \to x_2$, $x_5 \to x_3$, $x_4 \to x_4$, and $x_2 \to x_5$.

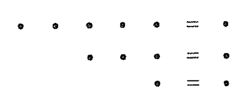


Figure 2. Was this case not covered?

The case where $m \geq n$ can be studied in a similar manner, and we can draw the following general conclusions.

THEOREM 8.3.2 *Existence/Uniqueness for Linear Systems*
If $m < n$, the system (1) can be consistent or inconsistent. If it is consistent it cannot have a unique solution; it will have a p-parameter family of solutions, where $n - m \leq p \leq n$. If $m \geq n$, (1) can be consistent or inconsistent. If it is

*The method is attributed to *Karl Friedrich Gauss* (1777–1855), who is generally regarded as the foremost mathematician of the nineteenth century and often referred to as the "prince of mathematicians."

consistent it can have a unique solution or a p-parameter family of solutions, where $1 \leq p \leq n$.

The next theorem follows immediately from Theorem 8.3.2, but we state it separately for emphasis.

THEOREM 8.3.3 *Existence / Uniqueness for Linear Systems*
Every system (1) necessarily admits no solution, a unique solution, or an infinity of solutions.

Observe that a system (1) is inconsistent only if, in its Gauss-eliminated form, one or more of the equations is of the form zero equal to a nonzero number. But that can never happen if every c_j in (1) is zero, that is, if (1) is **homogeneous**.

THEOREM 8.3.4 *Existence / Uniqueness for Homogeneous Systems*
Every homogeneous linear system of m equations in n unknowns is consistent. Either it admits the unique trivial solution or else it admits an infinity of nontrivial solutions in addition to the trivial solution. If $m < n$, then there is an infinity of nontrivial solutions in addition to the trivial solution.

In summary, not only did the method of Gauss elimination provide us with an efficient and systematic solution procedure, it also led us to important results regarding the existence and uniqueness of solutions.

8.3.3. Matrix notation. In applying Gauss elimination, we quickly discover that writing the variables $x_1, \ldots, x_n$ over and over is inefficient, and even tends to upstage the more central role of the a_{ij}'s and c_j's. It is therefore preferable to omit the x_j's altogether and to work directly with the rectangular array

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & c_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & c_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & c_m \end{bmatrix}, \tag{20}$$

known as the **augmented matrix** of the system (1), that is, the **coefficient matrix**

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \tag{21}$$

augmented by the column of c_j's. By *matrix* we simply mean a rectangular array of numbers, called **elements**; it is customary to enclose the elements between parentheses to emphasize that the entire matrix is regarded as a single entity. A horizontal line of elements is called a **row**, and a vertical line is called a **column**. Counting rows from the top, and columns from the left,

$$a_{21} \quad a_{22} \quad \cdots \quad a_{2n} \quad c_2 \qquad \text{and} \qquad \begin{matrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{matrix},$$

say, are the second row and $(n+1)$th column, respectively, of the augmented matrix (20).

In terms of the abbreviated matrix notation, the calculation in Example 1 would look like this.

Original system:

$$\begin{bmatrix} 1 & 1 & -1 & 1 \\ 3 & 1 & 1 & 9 \\ 1 & -1 & 4 & 8 \end{bmatrix}.$$

Add -3 times first row to second row, and add -1 times first row to third row:

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

Add -1 times second row to third row, and multiply second row by $-\frac{1}{2}$:

$$\begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \tag{22}$$

Thus, corresponding to the so-called elementary equation operations on members of a system of linear equations there are **elementary row operations** on the augmented matrix, as follows:

1. Addition of a multiple of one row to another:
   Symbolically: (jth row) $\rightarrow$ (jth row) + α(kth row)

2. Multiplication of a row by a nonzero constant:
   Symbolically: (jth row) $\rightarrow$ α(jth row)

3. Interchange of two rows:
   Symbolically: (jth row) $\leftrightarrow$ (kth row)

And we say that two matrices are **row equivalent** if one can be obtained from the other by finitely many elementary row operations.

8.3.4. Gauss–Jordan reduction. With the Gauss elimination completed, the remaining steps consist of back substitution. In fact, those steps are elementary row operations as well. The difference is that whereas in the Gauss elimination we proceed from the top down, in the back substitution we proceed from the bottom up.

EXAMPLE 4. To illustrate, let us return to Example 1 and pick up at the end of the Gauss elimination, with (5), and complete the back substitution steps using elementary row operations. In matrix format, we begin with

$$\begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \tag{23}$$

Keeping the bottom row intact, add 2 times that row to the second, and add 1 times that row to the first:

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \tag{24}$$

Now keeping the bottom two rows intact, add -1 times the second row to the first:

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \tag{25}$$

which is the solution: $x_1 = 3, x_2 = -1, x_3 = 1$ as obtained in Example 1. ∎

The entire process, of Gauss elimination plus back substitution, is known as **Gauss–Jordan reduction**, after Gauss and *Wilhelm Jordan* (1842–1899). The final result is an augmented matrix in **reduced row-echelon form**. That is:

1. In each row not made up entirely of zeros, the first nonzero element is a 1, a so-called **leading 1**.
2. In any two consecutive rows not made up entirely of zeros, the leading 1 in the lower row is to the right of the leading 1 in the upper row.
3. If a column contains a leading 1, every other element in that column is a zero.
4. All rows made up entirely of zeros are grouped together at the bottom of the matrix.

For instance, (25) is in reduced row-echelon form, as is the final matrix in the next example.

EXAMPLE 5. Let us return to Example 3 and finish the Gauss–Jordan reduction, beginning with (18):

$$\begin{bmatrix} 1 & -3 & 0 & -4 & -2 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & 2 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 1 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & \frac{3}{2} & 2 & 1 & \frac{5}{2} & 0 \\ 0 & 1 & \frac{1}{2} & 2 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 1 & -5 \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & 0 & \frac{3}{2} & 2 & 0 & \frac{3}{2} & 3 \\ 0 & 1 & \frac{1}{2} & 2 & 0 & -\frac{1}{2} & 3 \\ 0 & 0 & 0 & 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 1 & -5 \end{bmatrix} \to \begin{bmatrix} \mathbf{1} & 0 & \frac{3}{2} & 2 & 0 & 0 & \frac{21}{2} \\ 0 & \mathbf{1} & \frac{1}{2} & 2 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & -5 \end{bmatrix}.$$

The last augmented matrix is in reduced row-echelon form. The four leading 1's are displayed in **bold type**, and we see that, as a result of the back substitution steps, only 0's are to be found above each leading 1. The final augmented matrix once again gives the solution (19). ▮

8.3.5. Pivoting. Recall that the first step in the Gauss elimination of the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= c_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= c_2, \\ &\ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= c_m, \end{aligned} \tag{26}$$

is to subtract a_{21}/a_{11} times the first equation from the second, a_{31}/a_{11} times the first equation from the third, and so on, while keeping the first equation intact. The first equation is called the **pivot equation** (or, the first row is the pivot row if one is using the matrix format), and a_{11} is called the **pivot**. That step produces an indented system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= c_1, \\ a'_{22}x_2 + \cdots + a'_{2n}x_n &= c'_2, \\ &\ \ \vdots \\ a'_{m2}x_2 + \cdots + a'_{mn}x_n &= c'_m. \end{aligned} \tag{27}$$

Next, we keep the first *two* equations intact and use the second equation as the new pivot equation to indent the third through mth equations, and so on.

Naturally, we need each pivot to be nonzero. For instance, we need $a_{11} \neq 0$ for $a_{21}/a_{11}, a_{31}/a_{11}, \ldots$ to be defined. If a pivot is zero, interchange that equation with any one below it, such as the next equation or last equation (as we did in Example 3), until a nonzero pivot is available. Such interchange of equations is called **partial pivoting**. If a pivot is zero we have no choice but to use partial partial pivoting, but in practice even a nonzero pivot should be rejected if it is "very small," since the smaller it is the more susceptible is the calculation to the adverse effect of machine roundoff error (see Exercise 13). To be as safe as possible, one can choose the pivot equation as the one with the largest leading coefficient (relative to the other coefficients in the equation).

Closure. Beginning with a system of coupled linear algebraic equations, one can use a sequence of elementary operations to minimize the coupling between the equations while leaving the solution set intact. Besides putting forward the important method of Gauss elimination, which is used heavily in the following chapters, we used that method to establish several important theoretical results regarding the existence and uniqueness of solutions.

The Gauss elimination and Gauss–Jordan reduction discussions lead naturally to a convenient, and equivalent, formulation in matrix notation. We will return to the concept of matrices in Chapter 10, and develop it in detail.

Computer software. Chapters 8–12 cover the domain known as linear algebra. A great many calculations in linear algebra can be carried out using computer algebra systems. In *Maple*, for instance, a great many commands ("functions") are contained within the **linalg** package. A listing of these commands can be obtained by entering ?linalg. That list includes the **linsolve** command, which can be used to solve a system of m linear algebraic equations in n unknowns. To access linsolve (or any other command within the linalg package), first enter with(linalg). Then, linsolve(A,b) solves (1) for $x_1, \ldots, x_n$, where A is the coefficient matrix and b is the column of c_j's. For instance, the system

$$\begin{aligned} x_1 - x_2 + 2x_3 - 3x_4 &= 4, \\ x_1 + 2x_2 - x_3 + 3x_4 &= 1 \end{aligned} \tag{28}$$

admits the two-parameter family of solutions

$$x_4 = \alpha_1, \quad x_3 = \alpha_2, \quad x_2 = -1 - 2\alpha_1 + \alpha_2, \quad x_1 = 3 + \alpha_1 - \alpha_2, \tag{29}$$

where α_1, α_2 are arbitrary. To solve (28) using *Maple*, enter

with(linalg):

then return and enter

linsolve(array([[1, −1, 2, −3], [1, 2, −1, 3]]), array([4, 1]));

and return. The output is

$$[-\_t_1 + \_t_2 + 3, \ \ \_t_1 - 2\_t_2 - 1, \ \ \_t_1, \ \ \_t_2]$$

where the entries are $x_1, \ldots, x_4$ and where $\_t_1$ and $\_t_2$ are arbitrary constants. With $\_t_1 = \alpha_2$ and $\_t_2 = \alpha_1$, this result is the same as (29). If you prefer, you could use the sequence

```
with(linalg):
A := array([[1, -1, 2, -3], [1, 2, -1, 3]]):
b := array([4, 1]):
linsolve(A, b);
```

instead. If the system is inconsistent, then either the output will be NULL, or there will be no output.

EXERCISES 8.3

1. Derive the solution set for each of the following systems using Gauss elimination and augmented matrix format. Document each step (e.g., 2nd row→ 2nd row + 5 times 1st row), and classify the result (e.g., unique solution, the system is inconsistent, 3-parameter family of solutions, etc.).

(a) $2x - 3y = 1$
$5x + y = 2$

(b) $2x + y = 0$
$3x - 2y = 0$

(c) $x + 2y = 4$

(d) $x - y + z = 1$
$2x - y - z = 8$

(e) $2x_1 - x_2 - x_3 - 5x_4 = 6$

(f) $2x_1 - x_2 - x_3 - 3x_4 = 0$
$x_1 - x_2 + 4x_3 = 2$

(g) $x + 2y + 3z = 4$
$5x + 6y + 7z = 8$
$9x + 10y + 11z = 12$

(h) $x_1 + x_2 - 2x_3 = 3$
$x_1 - x_2 - 3x_3 = 1$
$x_1 - 3x_2 - 4x_3 = -1$

(i) $2x_1 - x_2 = 6$
$3x_1 + 2x_2 = 4$
$x_1 + 10x_2 = -12$
$6x_1 + 11x_2 = -2$

(j) $x_1 - x_2 + 2x_3 + x_4 = -1$
$2x_1 + x_2 + x_3 - x_4 = 4$
$x_1 + 2x_2 - x_3 - 2x_4 = 5$
$x_1 + x_3 = 1$

(k) $x_1 + x_3 = 1$
$x_1 + 2x_2 - x_3 - 2x_4 = 5$
$x_1 - x_2 + 2x_3 + x_4 = 0$
$2x_1 + x_2 + x_3 - x_4 = 4$

(l) $x_3 + x_4 = 2$
$4x_2 - x_3 + x_4 = 0$
$x_1 - x_2 + 2x_3 + x_4 = 4$

(m) $x + 2y + 3z = 5$
$2x + 3y + 4z = 8$
$3x + 4y + 5z = c$
$x + y = 2$
for $c = 10$, and again, for $c = 11$

(n) $2x + y + z = 10$
$3x + y - z = 6$
$x - 2y - 4z = -10$

(o) $2x_1 + x_2 = 1$
$x_1 + 2x_2 + x_3 = 1$
$x_2 + 2x_3 + x_4 = 1$
$x_3 + 2x_4 = 1$

(p) $2x_1 + x_2 = 0$
$x_1 + 2x_2 + x_3 = -1$
$x_2 + 2x_3 = -4$

(q) $2x_1 + x_2 + x_4 + 2x_5 = 0$
$x_1 + x_2 - x_3 = 0$
$x_1 + x_2 + x_3 - 3x_4 + 2x_5 = 0$
$2x_1 + 2x_2 - x_3 + x_5 = 0$

2. (a)–(q) Same as Exercise 1 but using Gauss–Jordan reduction instead of Gauss elimination.

3. (a)–(q) Same as Exercise 1 but using computer software such as the *Maple* linsolve command.

4. Can 20 linear algebraic equations in 14 unknowns have a unique solution? Be inconsistent? Have a two-parameter family of solutions? Have a 14-parameter family of solutions? Have a 16-parameter family of solutions? Explain.

5. Let

$$\begin{aligned} a_1x_1 + a_2x_2 + a_3x_3 &= 0, \\ b_1x_1 + b_2x_2 + b_3x_3 &= 0 \end{aligned}$$

represent any two planes through the origin in a Cartesian x_1, x_2, x_3 space. For the case where the planes intersect along a line, show whether or not that line necessarily passes through the origin.

6. If possible, adapt the methods of this section to solve the following *nonlinear* systems. If it is *not* possible, say so.

(a)
$$\begin{aligned} x_1^2 + 2x_2^2 - x_3^2 &= 29 \\ x_1^2 + x_2^2 + x_3^2 &= 19 \\ 3x_1^2 + 4x_2^2 \quad &= 67 \end{aligned}$$

(b)
$$\begin{aligned} x + 3y &= 13 \\ \sin x + 2y &= 5 \end{aligned}$$

(c)
$$\begin{aligned} \sin x + \sin y \quad &= 1 \\ \sin x - \sin y + 4\cos z &= 1.2 \\ \sin x + \sin y + 2\cos z &= 1.6 \end{aligned}$$

where $-\pi/2 \le x \le \pi/2$, $-\pi/2 \le y \le \pi/2$, and $0 \le z \le 2\pi$.

7. For what values of the parameter λ do the following homogeneous (do you agree that they are homogeneous?) systems admit *non*trivial solutions? Find the nontrivial solutions corresponding to each such λ.

(a)
$$\begin{aligned} 2x + y &= \lambda x \\ x + 2y &= \lambda y \end{aligned}$$

(b)
$$\begin{aligned} 2x - y &= \lambda x \\ -x + 2y &= \lambda y \end{aligned}$$

(c)
$$\begin{aligned} x - 2y &= \lambda x \\ 4x - 8y &= \lambda y \end{aligned}$$

(d)
$$\begin{aligned} z &= \lambda x \\ z &= \lambda y \\ x + y + z &= \lambda z \end{aligned}$$

(e)
$$\begin{aligned} x + y + z &= \lambda x \\ y + z &= \lambda y \\ 2z &= \lambda z \end{aligned}$$

(f)
$$\begin{aligned} 2x + y + z &= \lambda x \\ x + 2y + z &= \lambda y \\ x + y + 2z &= \lambda z \end{aligned}$$

8. Evaluate these excerpts from examination papers.

(a) "Given the system

$$\begin{aligned} x_1 - 2x_2 &= 0, \\ 2x_1 - 4x_2 &= 0, \end{aligned}$$

add -2 times the first equation to the second and add $-\frac{1}{2}$ times the second equation to the first. By these Gauss elimination steps we obtain the equivalent system $0 = 0$ and $0 = 0$, and hence the two-parameter family of solutions $x_1 = \alpha_1$ (arbitrary), $x_2 = \alpha_2$ (arbitrary)."

(b) "Given the system

$$\begin{aligned} x_1 + x_2 - 4x_3 &= 0, \\ 2x_1 - x_2 + x_3 &= 0, \end{aligned}$$

since both left-hand sides equal zero, they must equal each other. Hence we have the equation

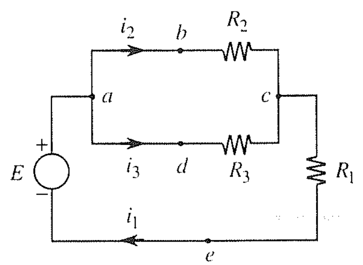
$$x_1 + x_2 - 4x_3 = 2x_1 - x_2 + x_3,$$

which equation is equivalent to the original system."

9. Make up an example of an inconsistent linear algebraic system of equations, with

(a) $m = 2, n = 4$ (b) $m = 1, n = 4$

10. (*Physical example of nonexistence and nonuniqueness; DC circuit*) Kirchoff's current and voltage laws were given in Section 2.3.1. If we apply those laws to the *DC* circuit shown,



we obtain the equations

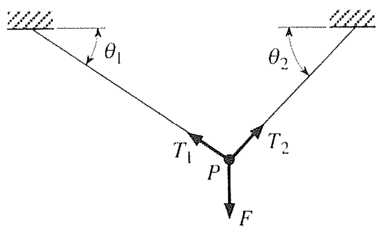
$$\begin{aligned} i_1 - i_2 - i_3 &= 0, && \text{(junction } a) \\ i_1 - i_2 - i_3 &= 0, && \text{(junction } c) \\ R_2 i_2 - R_3 i_3 &= 0, && \text{(loop } abcda) \\ R_1 i_1 + R_2 i_2 &= E, && \text{(loop } abcea) \\ R_1 i_1 + R_3 i_3 &= E, && \text{(loop } adcea) \end{aligned} \tag{10.1}$$

where i_1, i_2, i_3 are the three currents (measured as positive in the direction assumed in the figure), R_1, R_2, R_3 are the resistances of the three resistors, and E is the voltage rise (from e to a) induced by the battery or other source. [Evidently, we did not need to apply the current law to both junctions since the resulting equations are identical. Similarly, it may be that not all of the loop equations are needed. But rather than try to decide which of equations (10.1) to keep and which to discard, let us merely keep them all.] We now state the problem: Obtain the solution set of equations (10.1) by Gauss elimination. If there is no solution, or if there is a nonunique solution, explain that result in physical terms. Take

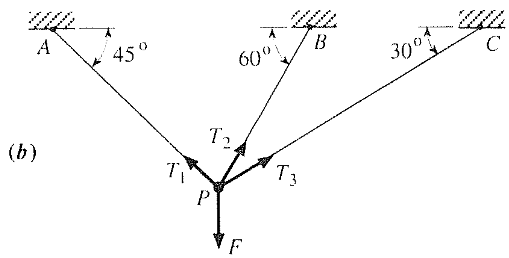
(a) $R_1 = R_2 = R_3 \equiv R \quad (R \neq 0)$
(b) $R_1 = R_2 \equiv R, \; R_2 = 2R \quad (R \neq 0)$
(c) $R_1 \equiv R, \; R_2 = R_3 = 2R \quad (R \neq 0)$

(d) $R_1 \equiv R,\ R_2 = 4R,\ R_3 = 6R \quad (R \neq 0)$
(e) $R_2 \equiv R,\ R_1 = R_3 = 0$
(f) $R_1 \equiv R,\ R_2 = R_3 = 0 \quad (R \neq 0)$
(g) $R_1 = R_2 = R_3 = 0$

11. (*Physical example of nonexistence and nonuniqueness; statically indeterminate structures*) (a) Consider the static equilibrium of the system shown, consisting of two weightless



cables connected at P, at which point a vertical load F is applied. Requiring an equilibrium of vertical force components, and horizontal force components too, derive two linear algebraic equations on the unknown tensions T_1 and T_2. Are there any combinations of angles θ_1 and θ_2 (where $0 \le \theta_1 \le \frac{\pi}{2}$ and $0 \le \theta_2 \le \frac{\pi}{2}$) such that there is either no solution or a nonunique solution? Explain.
(b) This time let there be three cables at angles of 45°, 60°, and 30° as shown. Again, requiring an equilibrium of vertical and



horizontal forces at P, derive two linear algebraic equations on the unknown tensions T_1, T_2, T_3. Show that the equations are consistent so there is a nonunique solution. NOTE: We say that such a structure is **statically indeterminate** because the forces in it cannot be determined from the laws of statics alone. What information needs to be added if we are to complete the evaluation of T_1, T_2, T_3? What is needed is information about the relative stiffness of the cables. We pursue this to a conclusion in (c), below.
(c) [*Completion of part (b)*] Before the load F is applied, locate an x, y Cartesian coordinate system at P. Let P be 1 foot below the "ceiling" so the coordinates of A, B, C are $(-1, 1)$, $(1/\sqrt{3}, 1)$, and $(\sqrt{3}, 1)$, respectively. Now apply the load F.

The point P will move to a point (x, y), and we assume that the cables are stiff enough so that x and y are very small: $|x| \ll 1$ and $|y| \ll 1$. Let the cables obey Hooke's law: $T_1 = k_1\delta_1$, $T_2 = k_2\delta_2$, and $T_3 = k_3\delta_3$, where δ_j is the increase in length of the jth cable due to the tension T_j. Since P moves to (x, y), it follows that

$$\begin{aligned}\delta_1 &= \sqrt{(x+1)^2 + (y-1)^2} - \sqrt{2}\\ &= \sqrt{2 + 2(x-y) + (x^2+y^2)} - \sqrt{2}\\ &\approx \sqrt{2 + 2(x-y)} - \sqrt{2}\\ &= \sqrt{2}\,[1 + (x-y)]^{1/2} - \sqrt{2}\\ &\approx \sqrt{2}\left[1 + \frac{1}{2}(x-y)\right] - \sqrt{2} = \frac{1}{\sqrt{2}}(x-y).\end{aligned} \tag{11.1}$$

Explain each step in (11.1), and show, similarly, that

$$\delta_2 \approx -\frac{1}{2}x - \frac{\sqrt{3}}{2}y, \tag{11.2}$$

$$\delta_3 \approx -\frac{\sqrt{3}}{2}x - \frac{1}{2}y. \tag{11.3}$$

Thus,

$$\begin{aligned}T_1 &= k_1\delta_1 \approx \frac{k_1}{\sqrt{2}}(x-y),\\ T_2 &= k_2\delta_2 \approx -\frac{k_2}{2}(x + \sqrt{3}y),\\ T_3 &= k_3\delta_3 \approx -\frac{k_3}{2}(\sqrt{3}x + y).\end{aligned} \tag{11.4}$$

Putting (11.4) into the two equilibrium equations obtained in (b) then gives two equations in the unknown displacements x, y. Show that that system can be solved uniquely for x and y, and thus complete the solution for T_1, T_2, T_3.

12. (*Roundoff error difficulty due to small pivots*) To illustrate how small pivots can accentuate the effects of roundoff error, consider the system

$$\begin{aligned}0.005x_1 + 1.47x_2 &= 1.49,\\ 0.975x_1 + 2.32x_2 &= 6.22\end{aligned} \tag{12.1}$$

with exact solution $x_1 = 4$ and $x_2 = 1$. Suppose that our computer carries three significant figures and then rounds off. Using the first equation as our pivot equation, Gauss elimination gives

$$\begin{bmatrix} 0.005 & 1.47 & 1.49 \\ 0.975 & 2.32 & 6.22 \end{bmatrix} \rightarrow \begin{bmatrix} 0.005 & 1.47 & 1.49 \\ 0 & -285 & -284 \end{bmatrix}$$

so $x_2 = 284/285 = 0.996$ and $x_1 = [1.49 - (1.47)(0.996)]/0.005 = (1.49 - 1.46)/0.005 = 6$. Show that if we use partial pivoting and then use the first equation of the system

$$\begin{aligned} 0.975x_1 + 2.32x_2 &= 6.22, \\ 0.005x_1 + 1.47x_2 &= 1.49 \end{aligned} \tag{12.2}$$

as our pivot equation, we obtain the result $x_1 = 4.00$ and $x_2 = 1.00$ (which happens to be exactly correct).

13. (*Ill-conditioned systems*) Practically speaking, our numerical calculations are normally carried out on computers, be they hand-held calculators or large digital computers. Such machines carry only a finite number of significant figures and thus introduce *roundoff error* into most calculations. One might expect (or hope) that such slight deviations will lead to answers that are only slightly in error. For example, the solution of

$$\begin{aligned} x + \quad y &= 2, \\ x - 1.014y &= 0 \end{aligned} \tag{13.1}$$

is $x \approx 1.007$, $y \approx 0.993$, whereas the solution of the rounded-off version

$$\begin{aligned} x + \quad y &= 2, \\ x - 1.01y &= 0 \end{aligned}$$

is very much the same, namely $x \approx 1.005$, $y \approx 0.995$. In sharp contrast, the solutions of

$$\begin{aligned} x + \quad y &= 2, \\ x + 1.014y &= 0 \end{aligned} \tag{13.2}$$

and the rounded-off version

$$\begin{aligned} x + \quad y &= 2, \\ x + 1.01y &= 0, \end{aligned}$$

$x \approx 144.9$, $y \approx -142.9$ and $x = 202$, $y = -200$, respectively; (13.2) is an example of a so-called **ill-conditioned** system (ill-conditioned in the sense that small changes in the coefficients lead to large changes in the solution). Here, we ask the following: Explain why (13.2) is much more sensitive to roundoff than (13.1) by exploring the two cases graphically, that is, in the x, y plane.

Chapter 8 Review

This chapter deals with systems of linear algebraic equations, m equations in n unknowns, insofar as the existence and uniqueness of solutions and solution technique. We find that there are three possibilities: a unique solution, no solution, and an infinity of solutions. If one or more solutions exist then the system is said to be consistent, if there are no solutions then it is inconsistent.

The key, in assessing existence/uniqueness as well as in finding solutions, is provided by elementary operations because they enable us to manipulate the system so as to reduce the coupling to a minimum, while at the same time keeping the solution set intact.

The method of Gauss elimination is introduced, as a systematic solution procedure based upon the three elementary operations, and it is shown that the subsequent back substitution steps amount to elementary operations as well. The entire process, Gauss elimination followed by the back substitution, is known as Gauss–Jordan reduction. Realize that the latter is a solution method, or algorithm, not a formula for the solution. Explicit solution formulas are developed, but not until Chapter 10.

We find that the process of Gauss elimination and Gauss–Jordan reduction are expressed most conveniently in matrix notation, although that notation is not essential to the method. In subsequent chapters the matrix approach is developed