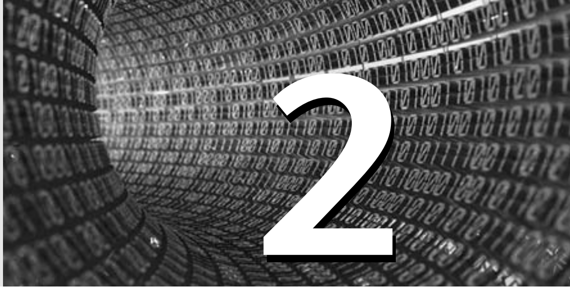

Simultaneous Linear Equations

2.1 Linear Systems

Systems of simultaneous equations appear frequently in engineering and scientific problems. Because of their importance and because they lend themselves to matrix analysis, we devote this entire chapter to their solutions.

We are interested in systems of the form

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1,\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2,\\
&\vdots\\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m.
\end{aligned} \tag{1}$$

We assume that the coefficients a_{ij} $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$ and the quantities b_i $(i = 1, 2, \ldots, m)$ are all known scalars. The quantities $x_1, x_2, \ldots, x_n$ represent unknowns.

Definition 1 A *solution* to (1) is a set of n scalars $x_1, x_2, \ldots, x_n$ that when substituted into (1) satisfies the given equations (that is, the equalities are valid).

System (1) is a generalization of systems considered earlier in that m can differ from n. If $m > n$, the system has more equations than unknowns. If $m < n$, the system has more unknowns than equations. If $m = n$, the system has as many unknowns as equations. In any case, the methods of Section 1.3 may be used to convert (1) into the matrix form

$$\mathbf{Ax} = \mathbf{b}, \tag{2}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Thus, if $m \neq n$, $\mathbf{A}$ will be rectangular and the dimensions of $\mathbf{x}$ and $\mathbf{b}$ will be different.

Example 1 Convert the following system to matrix form:

$$\begin{aligned} x + 2y - z + w &= 4, \\ x + 3y + 2z + 4w &= 9. \end{aligned}$$

Solution

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 1 & 3 & 2 & 4 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}. \quad \blacksquare$$

Example 2 Convert the following system to matrix form:

$$\begin{aligned} x - 2y &= -9, \\ 4x + y &= 9, \\ 2x + y &= 7, \\ x - y &= -1. \end{aligned}$$

Solution

$$\mathbf{A} = \begin{bmatrix} 1 & -2 \\ 4 & 1 \\ 2 & 1 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -9 \\ 9 \\ 7 \\ -1 \end{bmatrix}. \quad \blacksquare$$

A system of equations given by (1) or (2) can possess no solutions, exactly one solution, or more than one solution (note that by a solution to (2) we mean a vector $\mathbf{x}$ which satisfies the matrix equality (2)). Examples of such systems are

$$\begin{aligned} x + y &= 1, \\ x + y &= 2, \end{aligned} \tag{3}$$

$$\begin{aligned} x + y &= 1, \\ x - y &= 0, \end{aligned} \tag{4}$$

$$\begin{aligned} x + y &= 0, \\ 2x + 2y &= 0. \end{aligned} \tag{5}$$

Equation (3) has no solutions, (4) admits only the solution $x = y = \frac{1}{2}$, while (5) has solutions $x = -y$ for any value of y.

Definition 2 A system of simultaneous linear equations is *consistent* if it possesses at least one solution. If no solution exists, the system is *inconsistent*.

Equation (3) is an example of an inconsistent system, while (4) and (5) represent examples of consistent systems.

Definition 3 A system given by (2) is *homogeneous* if $\mathbf{b} = \mathbf{0}$ (the zero vector). If $\mathbf{b} \neq \mathbf{0}$ (at least one component of $\mathbf{b}$ differs from zero) the system is *nonhomogeneous*.

Equation (5) is an example of a homogeneous system.

Problems 2.1

In Problems 1 and 2, determine whether or not the proposed values of x, y, and z are solutions of the given systems.

1. $\begin{aligned} x + y + 2z &= 2, \\ x - y - 2z &= 0, \\ x + 2y + 2z &= 1. \end{aligned}$ (a) $x = 1, y = -3, z = 2$. (b) $x = 1, y = -1, z = 1$.

2. $\begin{aligned} x + 2y + 3z &= 6, \\ x - 3y + 2z &= 0, \\ 3x - 4y + 7z &= 6. \end{aligned}$ (a) $x = 1, y = 1, z = 1$. (b) $x = 2, y = 2, z = 0$. (c) $x = 14, y = 2, z = -4$.

3. Find a value for k such that $x = 1$, $y = 2$, and $z = k$ is a solution of the system

$$\begin{aligned} 2x + 2y + 4z &= 1, \\ 5x + y + 2z &= 5, \\ x - 3y - 2z &= -3. \end{aligned}$$

4. Find a value for k such that $x = 2$ and $y = k$ is a solution of the system

$$\begin{aligned} 3x + 5y &= 11, \\ 2x - 7y &= -3. \end{aligned}$$

5. Find a value for k such that $x = 2k$, $y = -k$, and $z = 0$ is a solution of the system

$$\begin{aligned} x + 2y + z &= 0, \\ -2x - 4y + 2z &= 0, \\ 3x - 6y - 4z &= 1. \end{aligned}$$

6. Find a value for k such that $x = 2k$, $y = -k$, and $z = 0$ is a solution of the system

$$\begin{aligned} x + 2y + 2z &= 0, \\ 2x - 4y + 2z &= 0, \\ -3x - 6y - 4z &= 0. \end{aligned}$$

7. Find a value for k such that $x = 2k$, $y = -k$, and $z = 0$ is a solution of the system

$$\begin{aligned} x + 2y + 2z &= 0, \\ 2x + 4y + 2z &= 0, \\ -3x - 6y - 4z &= 1. \end{aligned}$$

8. Put the system of equations given in Problem 4 into the matrix form $\mathbf{Ax} = \mathbf{b}$.

9. Put the system of equations given in Problem 1 into the matrix form $\mathbf{Ax} = \mathbf{b}$.

10. Put the system of equations given in Problem 2 into the matrix form $\mathbf{Ax} = \mathbf{b}$.

11. Put the system of equations given in Problem 6 into the matrix form $\mathbf{Ax} = \mathbf{b}$.

12. A manufacturer receives daily shipments of 70,000 springs and 45,000 pounds of stuffing for producing regular and support mattresses. Regular mattresses r require 50 springs and 30 pounds of stuffing; support mattresses s require 60 springs and 40 pounds of stuffing. The manufacturer wants to know how many mattresses of each type should be produced daily to utilize all available inventory. Show that this problem is equivalent to solving two equations in the two unknowns r and s.

13. A manufacturer produces desks and bookcases. Desks d require 5 hours of cutting time and 10 hours of assembling time. Bookcases b require 15 minutes of cutting time and one hour of assembling time. Each day, the manufacturer has available 200 hours for cutting and 500 hours for assembling. The manufacturer wants to know how many desks and bookcases should be scheduled for completion each day to utilize all available workpower. Show that this problem is equivalent to solving two equations in the two unknowns d and b.

14. A mining company has a contract to supply 70,000 tons of low-grade ore, 181,000 tons of medium-grade ore, and 41,000 tons of high-grade ore to a supplier. The company has three mines which it can work. Mine A produces 8000 tons of low-grade ore, 5000 tons of medium-grade ore, and 1000 tons of high-grade ore during each day of operation. Mine B produces 3000 tons of low-grade ore, 12,000 tons of medium-grade ore, and 3000 tons of high-grade ore for each day it is in operation. The figures for mine C are 1000, 10,000, and 2000, respectively. Show that the problem of determining how many days each mine must be operated to meet contractual demands without surplus is equivalent to solving a set of three equations in A, B, and C, where the unknowns denote the number of days each mine will be in operation.

15. A pet store has determined that each rabbit in its care should receive 80 units of protein, 200 units of carbohydrates, and 50 units of fat daily. The store carries four different types of feed that are appropriate for rabbits with the following compositions:

Feed	Protein units/oz	Carbohydrates units/oz	Fat units/oz
A	5	20	3
B	4	30	3
C	8	15	10
D	12	5	7

The store wants to determine a blend of these four feeds that will meet the daily requirements of the rabbits. Show that this problem is equivalent to solving three equations in the four unknowns A, B, C, and D, where each unknown denotes the number of ounces of that feed in the blend.

16. A small company computes its end-of-the-year bonus b as 5% of the net profit after city and state taxes have been paid. The city tax c is 2% of taxable income, while the state tax s is 3% of taxable income with credit allowed for the city tax as a pretax deduction. This year, taxable income was \$400,000. Show that b, c, and s are related by three simultaneous equations.

17. A gasoline producer has \$800,000 in fixed annual costs and incurs an additional variable cost of \$30 per barrel B of gasoline. The total cost C is the sum of the fixed and variable costs. The net sales S is computed on a wholesale price of \$40 per barrel. (a) Show that C, B, and S are related by two simultaneous equations. (b) Show that the problem of determining how many barrels must be produced to break even, that is, for net sales to equal cost, is equivalent to solving a system of three equations.

18. **(Leontief Closed Models)** A closed economic model involves a society in which all the goods and services produced by members of the society are consumed by those members. No goods and services are imported from without and none are exported. Such a system involves N members, each of whom produces goods or services and charges for their use. The problem is to determine the prices each member should charge for his or her labor so that everyone breaks even after one year. For simplicity, it is assumed that each member produces one unit per year.

Consider a simple closed system consisting of a farmer, a carpenter, and a weaver. The farmer produces one unit of food each year, the carpenter produces one unit of finished wood products each year, and the weaver produces one unit of clothing each year. Let p_1 denote the farmer's annual income (that is, the price she charges for her unit of food), let p_2 denote the carpenter's annual income (that is, the price he charges for his unit of finished wood products), and let p_3 denote the weaver's annual income. Assume on an annual basis that the farmer and the carpenter consume 40% each of the available food, while the weaver eats the remaining 20%. Assume that the carpenter uses 25% of the wood products he makes, while the farmer uses 30% and the weaver uses 45%. Assume further that the farmer uses 50% of the weaver's clothing while the carpenter uses 35% and the weaver consumes the remaining 15%. Show that a break-even equation for the farmer is

$$0.40p_1 + 0.30p_2 + 0.50p_3 = p_1,$$

while the break-even equation for the carpenter is

$$0.40p_1 + 0.25p_2 + 0.35p_3 = p_2.$$

What is the break-even equation for the weaver? Rewrite all three equations as a homogeneous system.

19. Paul, Jim, and Mary decide to help each other build houses. Paul will spend half his time on his own house and a quarter of his time on each of the houses of Jim and Mary. Jim will spend one third of his time on each of the three houses under construction. Mary will spend one sixth of her time on Paul's house, one third on Jim's house, and one half of her time on her own house. For tax purposes each must place a price on his or her labor, but they want to do so in a way that each will break even. Show that the process of determining break-even wages is a Leontief closed model comprised of three homogeneous equations.

20. Four third world countries each grow a different fruit for export and each uses the income from that fruit to pay for imports of the fruits from the other countries. Country A exports 20% of its fruit to country B, 30% to country C, 35% to country D, and uses the rest of its fruit for internal consumption. Country B exports 10% of its fruit to country A, 15% to country C, 35% to country D, and retains the rest for its own citizens. Country C does not export to country A; it divides its crop equally between countries B and D and its own people. Country D does not consume its own fruit. All of its fruit is for export with 15% going to country A, 40% to country B, and 45% to country C. Show that the problem of determining prices on the annual harvests of fruit so that each country breaks even is equivalent to solving four homogeneous equations in four unknowns.

21. **(Leontief Input–Output Models)** Consider an economy consisting of N sectors, with each producing goods or services unique to that sector. Let x_i denote the amount produced by the ith sector, measured in dollars. Thus x_i represents the dollar value of the supply of product i available in the economy. Assume that every sector in the economy has a demand for a proportion (which may be zero) of the output of every other sector. Thus, each sector j has a demand, measured in dollars, for the item produced in sector i. Let a_{ij} denote the proportion of item j's revenues that must be committed to the purchase of items from sector i in order for sector j to produce its goods or services. Assume also that there is an external demand, denoted by d_i and measured in dollars, for each item produced in the economy.

The problem is to determine how much of each item should be produced to meet external demand without creating a surplus of any item. Show that for a two sector economy, the solution to this problem is given by the supply/demand equations

$$\begin{aligned} &\underline{\text{supply}} \qquad \underline{\text{demand}} \\ &x_1 = a_{11}x_1 + a_{12}x_2 + d_1, \\ &x_2 = a_{21}x_1 + a_{22}x_2 + d_2. \end{aligned}$$

Show that this system is equivalent to the matrix equations

$$\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{d} \quad \text{and} \quad (\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{d}.$$

In this formulation, $\mathbf{A}$ is called the *consumption matrix* and $\mathbf{d}$ the *demand vector*.

22. Determine $\mathbf{A}$ and $\mathbf{d}$ in the previous problem if sector 1 must expend half of its revenues purchasing goods from its own sector and one third of its revenues purchasing goods from the other sector, while sector 2 must expend one quarter of its revenues purchasing items from sector 1 and requires nothing from itself. In addition, the demand for items from these two sectors are \$20,000 and \$30,000, respectively.

23. A small town has three primary industries, coal mining (sector 1), transportation (sector 2), and electricity (sector 3). Production of one dollar of coal requires the purchase of 10 cents of electricity and 20 cents of transportation. Production of one dollar of transportation requires the purchase of 2 cents of coal and 35 cents of electricity. Production of one unit of electricity requires the purchase of 10 cents of electricity, 50 cents of coal, and 30 cents of transportation. The town has external contracts for \$50,000 of coal, \$80,000 of transportation, and \$30,000 units of electricity. Show that the problem of determining how much coal, electricity, and transportation is required to supply the external demand without a surplus is equivalent to solving a Leontief input–output model. What are $\mathbf{A}$ and $\mathbf{d}$?

24. An economy consists of four sectors: energy, tourism, transportation, and construction. Each dollar of income from energy requires the expenditure of 20 cents on energy costs, 10 cents on transportation, and 30 cents on construction. Each dollar of income gotten by the tourism sector requires the expenditure of 20 cents on tourism (primarily in the form of complimentary facilities for favored customers), 15 cents on energy, 5 cents on transportation, and 30 cents on construction. Each dollar of income from transportation requires the expenditure of 40 cents on energy and 10 cents on construction; while each dollar of income from construction requires the expenditure of 5 cents on construction, 25 cents on energy, and 10 cents on transportation. The only external demand is for tourism, and this amounts to $5 million dollars a year. Show that the problem of determining how much energy, tourism, transportation, and construction is required to supply the external demand without a surplus is equivalent to solving a Leontief input–output model. What are $\mathbf{A}$ and $\mathbf{d}$?

25. A constraint is often imposed on each column of the consumption matrix of a Leontief input–output model, that the sum of the elements in each column be less than unity. Show that this guarantees that each sector in the economy is profitable.

2.2 Solutions by Substitution

Most readers have probably encountered simultaneous equations in high school algebra. At that time, matrices were not available; hence other methods were developed to solve these systems, in particular, the method of substitution. We review this method in this section. In the next section, we develop its matrix equivalent, which is slightly more efficient and, more importantly, better suited for computer implementations.

Consider the system given by (1):

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1,\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2,\\
&\ \ \vdots\\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}$$

The method of substitution is the following: take the first equation and solve for x_1 in terms of $x_2, x_3, \ldots, x_n$ and then substitute this value of x_1 into all the other equations, thus eliminating it from those equations. (If x_1 does not appear in the first equation, rearrange the equations so that it does. For example, one might have to interchange the order of the first and second equations.) This new set of equations is called the *first derived set*. Working with the first derived set, solve the second equation for x_2 in terms of $x_3, x_4, \ldots, x_n$ and then substitute this value of x_2 into the third, fourth, etc. equations, thus eliminating it. This new set is the second derived set. This process is kept up until the following set of equations is obtained:

$$\begin{aligned}
x_1 &= c_{12}x_2 + c_{13}x_3 + c_{14}x_4 + \cdots + c_{1n}x_n + d_1, \\
x_2 &= \qquad\quad c_{23}x_3 + c_{24}x_4 + \cdots + c_{2n}x_n + d_2, \\
x_3 &= \qquad\qquad\qquad\quad c_{34}x_4 + \cdots + c_{3n}x_n + d_3, \\
&\;\vdots \\
x_m &= \qquad\qquad c_{m,m+1}x_{m+1} + \cdots + c_{mn}x_n + d_m,
\end{aligned} \tag{6}$$

where the c_{ij}'s and the d_i's are some combination of the original a_{ij}'s and b_i's. System (6) can be quickly solved by back substitution.

Example 1 Use the method of substitution to solve the system

$$\begin{aligned}
r + 2s + t &= 3, \\
2r + 3s - t &= -6, \\
3r - 2s - 4t &= -2.
\end{aligned}$$

Solution By solving the first equation for r and then substituting it into the second and third equations, we obtain the first derived set

$$\begin{aligned}
r &= 3 - 2s - t, \\
-s - 3t &= -12, \\
-8s - 7t &= -11.
\end{aligned}$$

By solving the second equation for s and then substituting it into the third equation, we obtain the second derived set

$$\begin{aligned}
r &= 3 - 2s - t, \\
s &= 12 - 3t, \\
17t &= 85.
\end{aligned}$$

By solving for t in the third equation and then substituting it into the remaining equations (of which there are none), we obtain the third derived set

$$\begin{aligned}
r &= 3 - 2s - t, \\
s &= 12 - 3t, \\
t &= 5.
\end{aligned}$$

Thus, the solution is $t = 5$, $s = -3$, $r = 4$. ■

Example 2 Use the method of substitution to solve the system

$$\begin{aligned} x + y + 3z &= -1, \\ 2x - 2y - z &= 1, \\ 5x + y + 8z &= -2. \end{aligned}$$

Solution The first derived set is

$$\begin{aligned} x &= -1 - y - 3z, \\ -4y - 7z &= 3, \\ -4y - 7z &= 3. \end{aligned}$$

The second derived set is

$$\begin{aligned} x &= -1 - y - 3z, \\ y &= -\frac{3}{4} - \frac{7}{4}z, \\ 0 &= 0. \end{aligned}$$

Since the third equation can not be solved for z, this is as far as we can go. Thus, since we can not obtain a unique value for z, the first and second equations will not yield a unique value for x and y. *Caution*: The third equation does *not* imply that $z = 0$. On the contrary, this equation says nothing at all about z, consequently z is completely arbitrary. The second equation gives y in terms of z. Substituting this value into the first equation, we obtain x in terms of z. The solution therefore is $x = -\frac{1}{4} - \frac{5}{4}z$ and $y = -\frac{3}{4} - \frac{7}{4}z$, z is arbitrary. Thus there are infinitely many solutions to the above system. However, once z is chosen, x and y are determined. If z is chosen to be -1, then $x = y = 1$, while if z is chosen to be 3, then $x = -4$, $y = -6$. The solutions can be expressed in the vector form

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} & -\frac{5}{4}z \\ -\frac{3}{4} & -\frac{7}{4}z \\ & z \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} \\ -\frac{3}{4} \\ 0 \end{bmatrix} + z \begin{bmatrix} -\frac{5}{4} \\ -\frac{7}{4} \\ 1 \end{bmatrix}. \quad \blacksquare$$

Example 3 Use the method of substitution to solve

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

Solution The first derived set is

$$\begin{aligned} a &= 1 - 2b + 3c - d, \\ 2b + 10c &= 6. \end{aligned}$$

The second derived set is

$$\begin{aligned} a &= 1 - 2b + 3c - d \\ b &= 3 - 5c \end{aligned}$$

Again, since there are no more equations, this is as far as we can go, and since there are no defining equations for c and d, these two unknowns must be arbitrary. Solving for a and b in terms of c and d, we obtain the solution $a = -5 + 13c - d$, $b = 3 - 5c$; c and d are arbitrary. The solutions can be expressed in the vector form

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} -5 + 13c - d \\ 3 - 5c \\ c \\ d \end{bmatrix} = \begin{bmatrix} -5 \\ 3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 13 \\ -5 \\ 1 \\ 0 \end{bmatrix} + d\begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Note that while c and d are arbitrary, once they are given a particular value, a and b are automatically determined. For example, if c is chosen as -1 and d as 4, a solution is $a = -22$, $b = 8$, $c = -1$, $d = 4$, while if c is chosen as 0 and d as -3, a solution is $a = -2$, $b = 3$, $c = 0$, $d = -3$. ■

Example 4 Use the method of substitution to solve the following system:

$$\begin{aligned} x + 3y &= 4, \\ 2x - y &= 1, \\ 3x + 2y &= 5, \\ 5x + 15y &= 20. \end{aligned}$$

Solution The first derived set is

$$\begin{aligned} x &= 4 - 3y, \\ -7y &= -7, \\ -7y &= -7, \\ 0 &= 0. \end{aligned}$$

The second derived set is

$$\begin{aligned} x &= 4 - 3y, \\ y &= 1, \\ 0 &= 0, \\ 0 &= 0. \end{aligned}$$

Thus, the solution is $y = 1$, $x = 1$, or in vector form

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \quad \blacksquare$$

Problems 2.2

Use the method of substitution to solve the following systems:

1. $$\begin{aligned} x + 2y - 2z &= -1, \\ 2x + y + z &= 5, \\ -x + y - z &= -2. \end{aligned}$$

2. $$\begin{aligned} x + y - z &= 0, \\ 3x + 2y + 4z &= 0. \end{aligned}$$

3. $$\begin{aligned} x + 3y &= 4, \\ 2x - y &= 1, \\ -2x - 6y &= -8, \\ 4x - 9y &= -5, \\ -6x + 3y &= -3. \end{aligned}$$

4. $$\begin{aligned} 4r - 3s + 2t &= 1, \\ r + s - 3t &= 4, \\ 5r - 2s - t &= 5. \end{aligned}$$

5. $$\begin{aligned} 2l - m + n - p &= 1, \\ l + 2m - n + 2p &= -1, \\ l - 3m + 2n - 3p &= 2. \end{aligned}$$

6. $$\begin{aligned} 2x + y - z &= 0, \\ x + 2y + z &= 0, \\ 3x - y + 2z &= 0. \end{aligned}$$

7. $$\begin{aligned} x + 2y - z &= 5, \\ 2x - y + 2z &= 1, \\ 2x + 2y - z &= 7, \\ x + 2y + z &= 3. \end{aligned}$$

8. $$\begin{aligned} x + 2y + z - 2w &= 1, \\ 2x + 2y - z - w &= 3, \\ 2x - 2y + 2z + 3w &= 3, \\ 3x + y - 2z - 3w &= 1. \end{aligned}$$

2.3 Gaussian Elimination

Although the method of substitution is straightforward, it is not the most efficient way to solve simultaneous equations, and it does not lend itself well to electronic computing. Computers have difficulty symbolically manipulating the unknowns in algebraic equations. A striking feature of the method of substitution, however, is that the unknowns remain unaltered throughout the process: x remains x, y remains y, z remains z. Only the coefficients of the unknowns and the numbers on the right side of the equations change from one derived set to the next. Thus, we can save a good deal of writing, and develop a useful representation for computer processing, if we direct our attention to just the numbers themselves.

Definition 1 Given the system $\mathbf{Ax} = \mathbf{b}$, the *augmented matrix*, designated by $\mathbf{A^b}$, is a matrix obtained from $\mathbf{A}$ by adding to it one extra column, namely $\mathbf{b}$.

Thus, if

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 7 \\ 8 \end{bmatrix},$$

then

$$\mathbf{A^b} = \begin{bmatrix} 1 & 2 & 3 & 7 \\ 4 & 5 & 6 & 8 \end{bmatrix},$$

while if

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix},$$

then

$$\mathbf{A^b} = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 4 & 5 & 6 & -2 \\ 7 & 8 & 9 & -3 \end{bmatrix}.$$

In particular, the system

$$\begin{aligned} x + y - 2z &= -3, \\ 2x + 5y + 3z &= 11, \\ -x + 3y + z &= 5. \end{aligned}$$

has the matrix representation

$$\begin{bmatrix} 1 & 1 & -2 \\ 2 & 5 & 3 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \begin{bmatrix} -3 \\ 11 \\ 5 \end{bmatrix},$$

with an augmented matrix of

$$\mathbf{A^b} = \begin{bmatrix} 1 & 1 & -2 & -3 \\ 2 & 5 & 3 & 11 \\ -1 & 3 & 1 & 5 \end{bmatrix}.$$

Example 1 Write the set of equations in $x, y,$ and z associated with the augmented matrix

$$\mathbf{A^b} = \begin{bmatrix} -2 & 1 & 3 & 8 \\ 0 & 4 & 5 & -3 \end{bmatrix}.$$

Solution

$$\begin{aligned} -2x + y + 3z &= 8, \\ 4y + 5z &= -3. \end{aligned} \qquad \blacksquare$$

A second striking feature to the method of substitution is that every derived set is different from the system that preceded it. The method continues creating new derived sets until it has one that is particularly easy to solve by back-substitution. Of course, there is no purpose in solving any derived set, regardless how easy it is, unless we are assured beforehand that it has the same solution as the original system. Three elementary operations that alter equations but do not change their solutions are:

(i) Interchange the positions of any two equations.

(ii) Multiply an equation by a nonzero scalar.

(iii) Add to one equation a scalar times another equation.

If we restate these operations in words appropriate to an augmented matrix, we obtain the *elementary row operations:*

(E1) Interchange any two rows in a matrix.

(E2) Multiply any row of a matrix by a nonzero scalar.

(E3) Add to one row of a matrix a scalar times another row of that same matrix.

Gaussian elimination is a matrix method for solving simultaneous linear equations. The augmented matrix for the system is created, and then it is transformed into a row-reduced matrix (see Section 1.4) using elementary row operations. This is most often accomplished by using operation (E3) with each diagonal element in a matrix to create zeros in all columns directly below it, beginning with the first column and moving successively through the matrix, column by column. The system of equations associated with a row-reduced matrix can be solved easily by back-substitution, if we solve each equation for the first unknown that appears in it. This is the unknown associated with the first nonzero element in each nonzero row of the final augmented matrix.

Example 2 Use Gaussian elimination to solve

$$\begin{aligned} x + 3y &= 4, \\ 2x - y &= 1, \\ 3x + 2y &= 5, \\ 5x + 15y &= 20. \end{aligned}$$

Solution The augmented matrix for this system is

$$\begin{bmatrix} 1 & 3 & 4 \\ 2 & -1 & 1 \\ 3 & 2 & 5 \\ 5 & 15 & 20 \end{bmatrix}.$$

Then,

$$\begin{bmatrix} 1 & 3 & 4 \\ 2 & -1 & 1 \\ 3 & 2 & 5 \\ 5 & 15 & 20 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 4 \\ 0 & -7 & -7 \\ 3 & 2 & 5 \\ 5 & 15 & 20 \end{bmatrix} \quad \left\{ \begin{array}{l} \text{by adding to the} \\ \text{second row } (-2) \text{ times} \\ \text{the first row} \end{array} \right.$$

$$\to \begin{bmatrix} 1 & 3 & 4 \\ 0 & -7 & -7 \\ 0 & -7 & -7 \\ 5 & 15 & 20 \end{bmatrix} \quad \left\{ \begin{array}{l} \text{by adding to the} \\ \text{third row } (-3) \text{ times} \\ \text{the first row} \end{array} \right.$$

$$\to \begin{bmatrix} 1 & 3 & 4 \\ 0 & -7 & -7 \\ 0 & -7 & -7 \\ 0 & 0 & 0 \end{bmatrix} \quad \left\{ \begin{array}{l} \text{by adding to the} \\ \text{fourth row } (-5) \text{ times} \\ \text{the first row} \end{array} \right.$$

$$\to \begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 1 \\ 0 & -7 & -7 \\ 0 & 0 & 0 \end{bmatrix} \quad \left\{ \begin{array}{l} \text{by multiplying the} \\ \text{second row by } \dfrac{-1}{7} \end{array} \right.$$

$$\to \begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad \left\{ \begin{array}{l} \text{by adding to the} \\ \text{second row } (7) \text{ times} \\ \text{the first row} \end{array} \right.$$

The system of equations associated with this last augmented matrix in row-reduced form is

$$\begin{aligned} x + 3y &= 4, \\ y &= 1, \\ 0 &= 0, \\ 0 &= 0. \end{aligned}$$

Solving the second equation for y and then the first equation for x, we obtain $x = 1$ and $y = 1$, which is also the solution to the original set of equations. Compare this solution with Example 4 of the previous section. ■

The notation $(\rightarrow)$ should be read "is transformed into"; an equality sign is not correct because the transformed matrix is not equal to the original one.

Example 3 Use Gaussian elimination to solve

$$\begin{aligned} r + 2s + t &= 3, \\ 2r + 3s - t &= -6, \\ 3r - 2s - 4t &= -2. \end{aligned}$$

Solution The augmented matrix for this system is

$$\begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 3 & -1 & -6 \\ 3 & -2 & -4 & -2 \end{bmatrix}.$$

Then,

$$\begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 3 & -1 & -6 \\ 3 & -2 & -4 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -3 & -12 \\ 3 & -2 & -4 & -2 \end{bmatrix} \left\{ \begin{array}{l} \text{by adding to the} \\ \text{second row } (-2) \text{ times} \\ \text{the first row} \end{array} \right.$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -3 & -12 \\ 0 & -8 & -7 & -11 \end{bmatrix} \left\{ \begin{array}{l} \text{by adding to the} \\ \text{third row } (-3) \text{ times} \\ \text{the first row} \end{array} \right.$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 12 \\ 0 & -8 & -7 & -11 \end{bmatrix} \left\{ \begin{array}{l} \text{by multiplying the} \\ \text{second row by } (-1) \end{array} \right.$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 12 \\ 0 & 0 & 17 & 85 \end{bmatrix} \quad \begin{cases} \text{by adding to the} \\ \text{third row (8) times} \\ \text{the second row} \end{cases}$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 12 \\ 0 & 0 & 1 & 5 \end{bmatrix}. \quad \begin{cases} \text{by multiplying the} \\ \text{third row by } \left(\frac{1}{17}\right) \end{cases}$$

The system of equations associated with this last augmented matrix in row-reduced form is

$$\begin{aligned} r + 2s + \ \ t &= \ \ 3, \\ s + 3t &= 12, \\ t &= \ \ 5. \end{aligned}$$

Solving the third equation for t, then the second equation for s, and, lastly, the first equation for r, we obtain $r = 4$, $s = -3$, and $t = 5$, which is also the solution to the original set of equations. Compare this solution with Example 1 of the previous section. ■

Whenever one element in a matrix is used to cancel another element to zero by elementary row operation (E3), the first element is called the *pivot*. In Example 3, we first used the element in the 1–1 position to cancel the element in the 2–1 position, and then to cancel the element in the 3–1 position. In both of these operations, the unity element in the 1–1 position was the pivot. Later, we used the unity element in the 2–2 position to cancel the element −8 in the 3–2 position; here, the 2–2 element was the pivot.

While transforming a matrix into row-reduced form, it is advisable to adhere to three basic principles:

- Completely transform one column to the required form before considering another column.
- Work on columns in order, from left to right.
- Never use an operation if it will change a zero in a previously transformed column.

As a consequence of this last principle, one never involves the ith row of a matrix in an elementary row operation after the ith column has been transformed into its required form. That is, once the first column has the proper form, no pivot element should ever again come from the first row; once the second column has the proper form, no pivot element should ever again come from the second row; and so on.

When an element we want to use as a pivot is itself zero, we interchange rows using operation (E1).

Example 4 Use Gaussian elimination to solve

$$\begin{aligned} 2c + 3d &= 4, \\ a \quad + 3c + \ d &= 2, \\ a + b + 2c \quad &= 0. \end{aligned}$$

Solution The augmented matrix is

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

Normally, we would use the element in the 1–1 position to cancel to zero the two elements directly below it, but we cannot because it is zero. To proceed with the reduction process, we must interchange the first row with either of the other two rows. The choice is arbitrary.

$$\begin{bmatrix} 0 & 0 & 2 & 3 & 4 \\ 1 & 0 & 3 & 1 & 2 \\ 1 & 1 & 2 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ 0 & 0 & 2 & 3 & 4 \\ 1 & 1 & 2 & 0 & 0 \end{bmatrix} \qquad \left\{\begin{array}{l}\text{by interchanging the} \\ \text{first row with the} \\ \text{second row}\end{array}\right.$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ 0 & 0 & 2 & 3 & 4 \\ 0 & 1 & -1 & -1 & -2 \end{bmatrix}. \qquad \left\{\begin{array}{l}\text{by adding to the} \\ \text{third row } (-1) \text{ times} \\ \text{the first row}\end{array}\right.$$

Next, we would like to use the element in the 2–2 position to cancel to zero the element in the 3–2 position, but we cannot because that prospective pivot is zero. We use elementary row operation (E1) once again. The transformation yields

$$\rightarrow \begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ 0 & 1 & -1 & -1 & -2 \\ 0 & 0 & 2 & 3 & 4 \end{bmatrix} \qquad \left\{\begin{array}{l}\text{by interchanging the} \\ \text{second row with the} \\ \text{third row}\end{array}\right.$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ 0 & 1 & -1 & -1 & -2 \\ 0 & 0 & 1 & 1.5 & 2 \end{bmatrix}. \qquad \left\{\begin{array}{l}\text{by multiplying the} \\ \text{third row by } (0.5)\end{array}\right.$$

The system of equations associated with this last augmented matrix in row-reduced form is

$$\begin{aligned} a \quad + 3c + \quad d &= \ \ 2, \\ b - \ c - \quad d &= -2, \\ c + 1.5d &= \ \ 2. \end{aligned}$$

We use the third equation to solve for c, the second equation to solve for b, and the first equation to solve for a, because these are the unknowns associated with the first nonzero element of each nonzero row in the final augmented matrix. We have no defining equation for d, so this unknown remains arbitrary. The solution is, $a = -4 + 3.5d$, $b = -0.5d$, $c = 2 - 1.5d$, and d arbitrary, or in vector form

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} -4 + 3.5d \\ -0.5d \\ 2 - 1.5d \\ d \end{bmatrix} = \begin{bmatrix} -4 \\ 0 \\ 2 \\ 0 \end{bmatrix} + \frac{d}{2}\begin{bmatrix} 7 \\ -1 \\ -3 \\ 2 \end{bmatrix}.$$

This is also the solution to the original set of equations. ■

The derived set of equations associated with a row-reduced, augmented matrix may contain an absurd equation, such as $0 = 1$. In such cases, we conclude that the derived set is inconsistent, because no values of the unknowns can simultaneously satisfy all the equations. In particular, it is impossible to choose values of the unknowns that will make the absurd equation true. Since the derived set has the same solutions as the original set, it follows that the original set of equations is also inconsistent.

Example 5 Use Gaussian elimination to solve

$$\begin{aligned} 2x + 4y + 3z &= 8, \\ 3x - 4y - 4z &= 3, \\ 5x \qquad - z &= 12. \end{aligned}$$

Solution The augmented matrix for this system is

$$\begin{bmatrix} 2 & 4 & 3 & 8 \\ 3 & -4 & -4 & 3 \\ 5 & 0 & -1 & 12 \end{bmatrix}.$$

Then,

$$\begin{bmatrix} 2 & 4 & 3 & 8 \\ 3 & -4 & -4 & 3 \\ 5 & 0 & -1 & 12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1.5 & 4 \\ 3 & -4 & -4 & 3 \\ 5 & 0 & -1 & 12 \end{bmatrix} \quad \left\{\begin{array}{l}\text{by multiplying the} \\ \text{first row by } \left(\frac{1}{2}\right)\end{array}\right.$$

$$\to \begin{bmatrix} 1 & 2 & 1.5 & 4 \\ 0 & -10 & -8.5 & -9 \\ 5 & 0 & -1 & 12 \end{bmatrix} \quad \left\{\begin{array}{l}\text{by adding to the} \\ \text{second row } (-3) \text{ times} \\ \text{the first row}\end{array}\right.$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1.5 & 4 \\ 0 & -10 & -8.5 & -9 \\ 0 & -10 & -8.5 & -8 \end{bmatrix} \quad \left\{ \begin{array}{l} \text{by adding to the} \\ \text{third row } (-5) \text{ times} \\ \text{the first row} \end{array} \right.$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1.5 & 4 \\ 0 & 1 & 0.85 & 0.9 \\ 0 & -10 & -8.5 & -8 \end{bmatrix} \quad \left\{ \begin{array}{l} \text{by multiplying the} \\ \text{second row by } \left(\frac{-1}{10}\right) \end{array} \right.$$

$$\rightarrow \begin{bmatrix} 1 & 2 & 1.5 & 4 \\ 0 & 1 & 0.85 & 0.9 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad \left\{ \begin{array}{l} \text{by adding to the} \\ \text{third row } (10) \text{ times} \\ \text{the second row} \end{array} \right.$$

The system of equations associated with this last augmented matrix in row-reduced form is

$$\begin{aligned} x + 2y + 1.5z &= 4, \\ y + 0.85z &= 0.9, \\ 0 &= 1. \end{aligned}$$

Since no values of x, y, and z can make this last equation true, this system, as well as the original one, has no solution. ■

Finally, we note that most matrices can be transformed into a variety of row-reduced forms. If a row-reduced matrix has two nonzero rows, then a different row-reduced matrix is easily constructed by adding to the first row any nonzero constant times the second row. The equations associated with both augmented matrices, however, will have identical solutions.

Problems 2.3

In Problems 1 through 5, construct augmented matrices for the given systems of equations:

1. $$\begin{aligned} x + 2y &= -3, \\ 3x + y &= 1. \end{aligned}$$

2. $$\begin{aligned} x + 2y - z &= -1, \\ 2x - 3y + 2z &= 4. \end{aligned}$$

3. $$\begin{aligned} a + 2b &= 5, \\ -3a + b &= 13, \\ 4a + 3b &= 0. \end{aligned}$$

4. $$\begin{aligned} 2r + 4s \phantom{{}+7t} &= 2, \\ 3r + 2s + t &= 8, \\ 5r - 3s + 7t &= 15. \end{aligned}$$

5. $$\begin{aligned} 2r + 3s - 4t &= 12, \\ 3r - 2s \phantom{{}-4t} &= -1, \\ 8r - s - 4t &= 10. \end{aligned}$$

In Problems 6 through 11, write the set of equations associated with the given augmented matrix and the specified variables.

6. $\mathbf{A^b} = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 8 \end{bmatrix}$ variables: x and y.

7. $\mathbf{A^b} = \begin{bmatrix} 1 & -2 & 3 & 10 \\ 0 & 1 & -5 & -3 \\ 0 & 0 & 1 & 4 \end{bmatrix}$ variables: x, y, and z.

8. $\mathbf{A^b} = \begin{bmatrix} 1 & -3 & 12 & 40 \\ 0 & 1 & -6 & -200 \\ 0 & 0 & 1 & 25 \end{bmatrix}$ variables: r, s, and t.

9. $\mathbf{A^b} = \begin{bmatrix} 1 & 3 & 0 & -8 \\ 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ variables: x, y, and z.

10. $\mathbf{A^b} = \begin{bmatrix} 1 & -7 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ variables: a, b, and c.

11. $\mathbf{A^b} = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ variables: u, v, and w.

12. Solve the system of equations defined in Problem 6.

13. Solve the system of equations defined in Problem 7.

14. Solve the system of equations defined in Problem 8.

15. Solve the system of equations defined in Problem 9.

16. Solve the system of equations defined in Problem 10.

17. Solve the system of equations defined in Problem 11.

In Problems 18 through 24, use elementary row operations to transform the given matrices into row-reduced form:

18. $\begin{bmatrix} 1 & -2 & 5 \\ -3 & 7 & 8 \end{bmatrix}$.

19. $\begin{bmatrix} 4 & 24 & 20 \\ 2 & 11 & -8 \end{bmatrix}$.

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

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

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

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

24. $\begin{bmatrix} 2 & 3 & 4 & 6 & 0 & 10 \\ -5 & -8 & 15 & 1 & 3 & 40 \\ 3 & 3 & 5 & 4 & 4 & 20 \end{bmatrix}.$

25. Solve Problem 1.

26. Solve Problem 2.

27. Solve Problem 3.

28. Solve Problem 4.

29. Solve Problem 5.

30. Use Gaussian elimination to solve Problem 1 of Section 2.2.

31. Use Gaussian elimination to solve Problem 2 of Section 2.2.

32. Use Gaussian elimination to solve Problem 3 of Section 2.2.

33. Use Gaussian elimination to solve Problem 4 of Section 2.2.

34. Use Gaussian elimination to solve Problem 5 of Section 2.2.

35. Determine a production schedule that satisfies the requirements of the manufacturer described in Problem 12 of Section 2.1.

36. Determine a production schedule that satisfies the requirements of the manufacturer described in Problem 13 of Section 2.1.

37. Determine a production schedule that satisfies the requirements of the manufacturer described in Problem 14 of Section 2.1.

38. Determine feed blends that satisfy the nutritional requirements of the pet store described in Problem 15 of Section 2.1.

39. Determine the bonus for the company described in Problem 16 of Section 2.1.

40. Determine the number of barrels of gasoline that the producer described in Problem 17 of Section 2.1 must manufacture to break even.

41. Determine the annual incomes of each sector of the Leontief closed model described in Problem 18 of Section 2.1.

42. Determine the wages of each person in the Leontief closed model described in Problem 19 of Section 2.1.

43. Determine the total sales revenue for each country of the Leontief closed model described in Problem 20 of Section 2.1.

44. Determine the production quotas for each sector of the economy described in Problem 22 of Section 2.1.

45. An *elementary matrix* is a square matrix **E** having the property that the product **EA** is the result of applying a single elementary row operation on the matrix **A**. Form a matrix **H** from the 4×4 identity matrix **I** by interchanging any two rows of **I**, and then compute the product **HA** for any 4×4 matrix **A** of your choosing. Is $\mathbf{H}$ an elementary matrix? How would one construct elementary matrices corresponding to operation (E1)?

46. Form a matrix $\mathbf{G}$ from the 4×4 identity matrix $\mathbf{I}$ by multiplying any one row of $\mathbf{I}$ by the number 5, and then compute the product $\mathbf{GA}$ for any 4×4 matrix $\mathbf{A}$ of your choosing. Is $\mathbf{G}$ an elementary matrix? How would one construct elementary matrices corresponding to operation (E2)?

47. Form a matrix $\mathbf{F}$ from the 4×4 identity matrix $\mathbf{I}$ by adding to one row of $\mathbf{I}$ five times another row of $\mathbf{I}$. Use any two rows of your choosing. Compute the product $\mathbf{FA}$ for any 4×4 matrix $\mathbf{A}$ of your choosing. Is $\mathbf{F}$ an elementary matrix? How would one construct elementary matrices corresponding to operation (E3)?

48. A solution procedure uniquely suited to matrix equations of the form $\mathbf{x} = \mathbf{Ax} + \mathbf{d}$ is iteration. A trial solution $\mathbf{x}^{(0)}$ is proposed, and then progressively better estimates $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \ldots$ for the solution are obtained iteratively from the formula

$$\mathbf{x}^{(i+1)} = \mathbf{Ax}^{(i)} + \mathbf{d}.$$

The iterations terminate when two successive estimates differ by less than a prespecified acceptable tolerance.

If the system comes from a Leontief input–output model, then a reasonable initialization is $\mathbf{x}^{(0)} = 2\mathbf{d}$. Apply this method to the system defined by Problem 22 of Section 2.1. Stop after two iterations.

49. Use the iteration method described in the previous problem to solve the system defined in Problem 23 of Section 2.1. In particular, find the first two iterations by hand calculations, and then use a computer to complete the iteration process.

50. Use the iteration method described in Problem 48 to solve the system defined in Problem 24 of Section 2.1. In particular, find the first two iterations by hand calculations, and then use a computer to complete the iteration process.

2.4 Pivoting Strategies

Gaussian elimination is often programmed for computer implementation. Since all computers round or truncate numbers to a finite number of digits (e.g., the fraction 1/3 might be stored as 0.33333, but never as the *infinite* decimal 0.333333...) roundoff error can be significant. A number of strategies have been developed to minimize the effects of such errors.

The most popular strategy is *partial pivoting*, which requires that a pivot element always be larger in absolute value than any element below it in the same column. This is accomplished by interchanging rows whenever necessary.

Example 1 Use partial pivoting with Gaussian elimination to solve the system

$$\begin{aligned} x + \;\; 2y + 4z &= 18, \\ 2x + 12y - 2z &= \;\; 9, \\ 5x + 26y + 5z &= 14. \end{aligned}$$

Solution The augmented matrix for this system is

$$\begin{bmatrix} 1 & 2 & 4 & 18 \\ 2 & 12 & -2 & 9 \\ 5 & 26 & 5 & 14 \end{bmatrix}.$$

Normally, the unity element in the 1–1 position would be the pivot. With partial pivoting, we compare this prospective pivot to all elements directly below it in the same column, and if any is larger in absolute value, as is the case here with the element 5 in the 3–1 position, we interchange rows to bring the largest element into the pivot position.

$$\begin{bmatrix} 1 & 2 & 4 & 18 \\ 2 & 12 & -2 & 9 \\ 5 & 26 & 5 & 14 \end{bmatrix} \to \begin{bmatrix} 5 & 26 & 5 & 14 \\ 2 & 12 & -2 & 9 \\ 1 & 2 & 4 & 18 \end{bmatrix}. \quad \left\{\begin{array}{l}\text{by interchanging the} \\ \text{first and third rows}\end{array}\right.$$

Then,

$$\to \begin{bmatrix} 1 & 5.2 & 1 & 2.8 \\ 2 & 12 & -2 & 9 \\ 1 & 2 & 4 & 18 \end{bmatrix} \quad \left\{\begin{array}{l}\text{by multiplying the} \\ \text{first row by } \frac{1}{5}\end{array}\right.$$

$$\to \begin{bmatrix} 1 & 5.2 & 1 & 2.8 \\ 0 & 1.6 & -4 & 3.4 \\ 1 & 2 & 4 & 18 \end{bmatrix} \quad \left\{\begin{array}{l}\text{by adding to the} \\ \text{second row } (-2) \text{ times} \\ \text{the first row}\end{array}\right.$$

$$\to \begin{bmatrix} 1 & 5.2 & 1 & 2.8 \\ 0 & 1.6 & -4 & 3.4 \\ 0 & -3.2 & 3 & 15.2 \end{bmatrix}. \quad \left\{\begin{array}{l}\text{by adding to the} \\ \text{third row } (-1) \text{ times} \\ \text{the first row}\end{array}\right.$$

The next pivot would normally be the element 1.6 in the 2–2 position. Before accepting it, however, we compare it to all elements directly below it in the same column. The largest element in absolute value is the element −3.2 in the 3–2 position. Therefore, we interchange rows to bring this larger element into the pivot position.

NOTE. We do not consider the element 5.2 in the 1–2 position, even though it is the largest element in its column. Comparisons are only made between a prospective pivot and all elements directly below it. Recall one of the three basic principles of row-reduction: never involve the first row of matrix in a row operation after the first column has been transformed into its required form.

$$\rightarrow \begin{bmatrix} 1 & 5.2 & 1 & 2.8 \\ 0 & -3.2 & 3 & 15.2 \\ 0 & 1.6 & -4 & 3.4 \end{bmatrix} \quad \begin{cases} \text{by interchanging the} \\ \text{second and third rows} \end{cases}$$

$$\rightarrow \begin{bmatrix} 1 & 5.2 & 1 & 2.8 \\ 0 & 1 & -0.9375 & -4.75 \\ 0 & 1.6 & -4 & 3.4 \end{bmatrix} \quad \begin{cases} \text{by multiplying the} \\ \text{second row by } \frac{-1}{3.2} \end{cases}$$

$$\rightarrow \begin{bmatrix} 1 & 5.2 & 1 & 2.8 \\ 0 & 1 & -0.9375 & -4.75 \\ 0 & 0 & -2.5 & 11 \end{bmatrix} \quad \begin{cases} \text{by adding to the} \\ \text{third row } (-1.6) \text{ times} \\ \text{the second row} \end{cases}$$

$$\rightarrow \begin{bmatrix} 1 & 5.2 & 1 & 2.8 \\ 0 & 1 & -0.9375 & -4.75 \\ 0 & 0 & 1 & -4.4 \end{bmatrix} \quad \begin{cases} \text{by multiplying the} \\ \text{third row by } \frac{-1}{2.5} \end{cases}$$

The new derived set of equations is

$$\begin{aligned} x + 5.2y + \qquad z &= 2.8, \\ y - 0.9375z &= -4.75, \\ z &= -4.4, \end{aligned}$$

which has as its solution $x = 53.35$, $y = -8.875$, and $z = -4.4$. ■

Scaled pivoting involves ratios. A prospective pivot is divided by the largest element in absolute value in its row, ignoring the last column. The result is compared to the ratios formed by dividing every element directly below the pivot by the largest element in absolute value in its respective row, again ignoring the last column. Of these, the element that yields the largest ratio in absolute value is designated as the pivot, and if that element is not already in the pivot position, then row interchanges are performed to move it there.

Example 2 Use scaled pivoting with Gaussian elimination to solve the system given in Example 1.

Solution The augmented matrix for this system is

$$\begin{bmatrix} 1 & 2 & 4 & 18 \\ 2 & 12 & -2 & 9 \\ 5 & 26 & 5 & 14 \end{bmatrix}.$$

Normally, we would use the element in the 1–1 position as the pivot. With scaled pivoting, however, we first compare ratios between elements in the first column to the largest elements in absolute value in each row, ignoring the last column. The ratios are

$$\frac{1}{4} = 0.25, \quad \frac{2}{12} = 0.1667, \quad \text{and} \quad \frac{5}{26} = 0.1923.$$

The largest ratio in absolute value corresponds to the unity element in the 1–1 position, so that element remains the pivot. Transforming the first column into reduced form, we obtain

$$\begin{bmatrix} 1 & 2 & 4 & 18 \\ 0 & 8 & -10 & -27 \\ 0 & 16 & -15 & -76 \end{bmatrix}.$$

Normally, the next pivot would be the element in the 2–2 position. Instead, we consider the ratios

$$\frac{8}{10} = 0.8 \quad \text{and} \quad \frac{16}{16} = 1,$$

which are obtained by dividing the pivot element and every element directly below it by the largest element in absolute value appearing in their respective rows, ignoring elements in the last column. The largest ratio in absolute value corresponds to the element 16 appearing in the 3–2 position. We move it into the pivot position by interchanging the second and third rows. The new matrix is

$$\begin{bmatrix} 1 & 2 & 4 & 18 \\ 0 & 16 & -15 & -76 \\ 0 & 8 & -10 & -27 \end{bmatrix}.$$

Completing the row-reduction transformation, we get

$$\begin{bmatrix} 1 & 2 & 4 & 18 \\ 0 & 1 & -0.9375 & -4.75 \\ 0 & 0 & 1 & -4.4 \end{bmatrix}.$$

The system of equations associated with this matrix is

$$\begin{aligned} x + 2y + 4z &= 18, \\ y - 0.9375z &= -4.75, \\ z &= -4.4. \end{aligned}$$

The solution is, as before, $x = 53.35$, $y = -8.875$, and $z = -4.4$. ■

Complete pivoting compares prospective pivots with all elements in the largest submatrix for which the prospective pivot is in the upper left position, ignoring the last column. If any element in this submatrix is larger in absolute value than the prospective pivot, both row and column interchanges are made to move this larger element into the pivot position. Because column interchanges rearrange the order of the unknowns, a book keeping method must be implemented to record all rearrangements. This is done by adding a new row, designated as row 0, to the matrix. The entries in the new row are initially the positive integers in ascending order, to denote that column 1 is associated with variable 1, column 2 with variable 2, and so on. This new top row is only affected by column interchanges; *none of the elementary row operations is applied to it.*

Example 3 Use complete pivoting with Gaussian elimination to solve the system given in Example 1.

Solution The augmented matrix for this system is

$$\left[\begin{array}{rrrr} 1 & 2 & 3 & \\ \hline 1 & 2 & 4 & 18 \\ 2 & 12 & -2 & 9 \\ 5 & 26 & 5 & 14 \end{array}\right].$$

Normally, we would use the element in the 1–1 position of the coefficient matrix $\mathbf{A}$ as the pivot. With complete pivoting, however, we first compare this prospective pivot to all elements in the submatrix shaded below. In this case, the element 26 is the largest, so we interchange rows and columns to bring it into the pivot position.

$$\left[\begin{array}{rrrr} 1 & 2 & 3 & \\ \hline 1 & 2 & 4 & 18 \\ 2 & 12 & -2 & 9 \\ 5 & 26 & 5 & 14 \end{array}\right] \to \left[\begin{array}{rrrr} 1 & 2 & 3 & \\ \hline 5 & 26 & 5 & 14 \\ 2 & 12 & -2 & 9 \\ 1 & 2 & 4 & 18 \end{array}\right] \quad \left\{\begin{array}{l}\text{by interchanging the}\\ \text{first and third rows}\end{array}\right.$$

$$\to \left[\begin{array}{rrrr} 2 & 1 & 3 & \\ \hline 26 & 5 & 5 & 14 \\ 12 & 2 & -2 & 9 \\ 2 & 1 & 4 & 18 \end{array}\right]. \quad \left\{\begin{array}{l}\text{by interchanging the}\\ \text{first and second columns}\end{array}\right.$$

Applying Gaussian elimination to the first column, we obtain

$$\left[\begin{array}{rrrr} 2 & 1 & 3 & \\ \hline 1 & 0.1923 & 0.1923 & 0.5385 \\ 0 & -0.3077 & -4.3077 & 2.5385 \\ 0 & 0.6154 & 3.6154 & 16.9231 \end{array}\right].$$

Normally, the next pivot would be -0.3077. Instead, we compare this number in absolute value to all the numbers in the submatrix shaded above. The largest such element in absolute value is -4.3077, which we move into the pivot position by interchanging the second and third column. The result is

$$\begin{array}{c} \begin{array}{cccc} 2 & 3 & 1 & \end{array} \\ \begin{bmatrix} 1 & 0.1923 & 0.1923 & 0.5385 \\ 0 & -4.3077 & -0.3077 & 2.5385 \\ 0 & 3.6154 & 0.6154 & 16.9231 \end{bmatrix} \end{array}.$$

Continuing with Gaussian elimination, we obtain the row-reduced matrix

$$\begin{array}{c} \begin{array}{cccc} 2 & 3 & 1 & \end{array} \\ \begin{bmatrix} 1 & 0.1923 & 0.1923 & 0.5385 \\ 0 & 1 & 0.0714 & -0.5893 \\ 0 & 0 & 1 & 53.35 \end{bmatrix} \end{array}.$$

The system associated with this matrix is

$$\begin{aligned} y + 0.1923z + 0.1923x &= 0.5385, \\ z + 0.0714x &= -0.5893, \\ x &= 53.35. \end{aligned}$$

Its solution is, $x = 53.35$, $y = -8.8749$, and $z = -4.3985$, which is within round-off error of the answers gotten previously. ■

Complete pivoting generally identifies a better pivot than scaled pivoting which, in turn, identifies a better pivot than partial pivoting. Nonetheless, partial pivoting is most often the strategy of choice. Pivoting strategies are used to avoid roundoff error. We do not need the best pivot; we only need to avoid bad pivots.

Problems 2.4

In Problems 1 through 6, determine the first pivot under (a) partial pivoting, (b) scaled pivoting, and (c) complete pivoting for given augmented matrices.

1. $\begin{bmatrix} 1 & 3 & 35 \\ 4 & 8 & 15 \end{bmatrix}$.

2. $\begin{bmatrix} 1 & -2 & -5 \\ 5 & 3 & 85 \end{bmatrix}$.

3. $\begin{bmatrix} 1 & 8 & 15 \\ 3 & -4 & 11 \end{bmatrix}$.

4. $\begin{bmatrix} -2 & 8 & -3 & 100 \\ 4 & 5 & 4 & 75 \\ -3 & -1 & 2 & 250 \end{bmatrix}$.

5. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$.

6. $\begin{bmatrix} 0 & 2 & 3 & 4 & 0 \\ 1 & 0.4 & 0.8 & 0.1 & 90 \\ 4 & 10 & 1 & 8 & 40 \end{bmatrix}$.

7. Solve Problem 3 of Section 2.3 using Gaussian elimination with each of the three pivoting strategies.

8. Solve Problem 4 of Section 2.3 using Gaussian elimination with each of the three pivoting strategies.

9. Solve Problem 5 of Section 2.3 using Gaussian elimination with each of the three pivoting strategies.

10. Computers internally store numbers in formats similar to the scientific notation 0, –E–, representing the number 0. –multiplied by the power of 10 signified by the digits following E. Therefore, 0.1234E06 is 123,400 while 0.9935E02 is 99.35. The number of digits between the decimal point and E is finite and fixed; it is the number of significant figures. Arithmetic operations in computers are performed in registers, which have twice the number of significant figures as storage locations.

Consider the system

$$0.00001x + y = 1.00001,$$

$$x + y = 2.$$

Show that when Gaussian elimination is implemented on this system by a computer limited to four significant figures, the result is $x = 0$ and $y = 1$, which is incorrect. Show further that the difficulty is resolved when partial pivoting is employed.

2.5 Linear Independence

We momentarily digress from our discussion of simultaneous equations to develop the concepts of linearly independent vectors and rank of a matrix, both of which will prove indispensable to us in the ensuing sections.

Definition 1 A vector $\mathbf{V}_1$ is a *linear combination* of the vectors $\mathbf{V}_2, \mathbf{V}_3, \ldots, \mathbf{V}_n$ if there exist scalars $d_2, d_3, \ldots, d_n$ such that

$$\mathbf{V}_1 = d_2\mathbf{V}_2 + d_3\mathbf{V}_3 + \cdots + d_n\mathbf{V}n.$$

Example 1 Show that $[1 \quad 2 \quad 3]$ is a linear combination of $[2 \quad 4 \quad 0]$ and $[0 \quad 0 \quad 1]$.

Solution $[1 \quad 2 \quad 3] = \frac{1}{2}[2 \quad 4 \quad 0] + 3[0 \quad 0 \quad 1]$. ■

Referring to Example 1, we could say that the row vector $[1 \quad 2 \quad 3]$ depends linearly on the other two vectors or, more generally, that the set of vectors $\{[1 \quad 2 \quad 3], [2 \quad 4 \quad 0], [0 \quad 0 \quad 1]\}$ is *linearly dependent*. Another way of expressing this dependence would be to say that there exist constants c_1, c_2, c_3 not all zero such that $c_1\ [1 \quad 2 \quad 3] + c_2\ [2 \quad 4 \quad 0] + c_3\ [0 \quad 0 \quad 1] = [0 \quad 0 \quad 0]$. Such a set would be $c_1 = -1, c_2 = \frac{1}{2}, c_3 = 3$. Note that the set $c_1 = c_2 = c_3 = 0$ is also a suitable set. The important fact about dependent sets, however, is that there exists a set of constants, *not all equal to zero*, that satisfies the equality.

Now consider the set given by $\mathbf{V}_1 = [1 \quad 0 \quad 0]\ \mathbf{V}_2 = [0 \quad 1 \quad 0]\ \mathbf{V}_3 = [0 \quad 0 \quad 1]$. It is easy to verify that no vector in this set is a linear combination of the other two. Thus, each vector is linearly independent of the other two or, more generally, the set of vectors is *linearly independent*. Another way of expressing this independence would be to say the only scalars that satisfy the equation $c_1[1 \quad 0 \quad 0] + c_2[0 \quad 1 \quad 0] + c_3[0 \quad 0 \quad 1] = [0 \quad 0 \quad 0]$ are $c_1 = c_2 = c_3 = 0$.

Definition 2 A set of vectors $\{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n\}$, of the same dimension, is *linearly dependent* if there exist scalars $c_1, c_2, \ldots, c_n$, not all zero, such that

$$c_1\mathbf{V}_1 + c_2\mathbf{V}_2 + c_3\mathbf{V}_3 + \cdots + c_n\mathbf{V}_n = \mathbf{0} \tag{7}$$

The vectors are *linearly independent* if the only set of scalars that satisfies (7) is the set $c_1 = c_2 = \cdots = c_n = 0$.

Therefore, to test whether or not a given set of vectors is linearly independent, first form the vector equation (7) and ask "What values for the c's satisfy this equation?" Clearly $c_1 = c_2 = \cdots = c_n = 0$ is a suitable set. If this is the only set of values that satisfies (7) then the vectors are linearly independent. If there exists a set of values that is not all zero, then the vectors are linearly dependent.

Note that it is not necessary for all the c's to be different from zero for a set of vectors to be linearly dependent. Consider the vectors $\mathbf{V}_1 = [1, 2], \mathbf{V}_2 = [1, 4], \mathbf{V}_3 = [2, 4]$. $c_1 = 2, c_2 = 0, c_3 = -1$ is a set of scalars, *not all zero,* such that $c_1\ \mathbf{V}_1 + c_2\ \mathbf{V}_2 + c_3\mathbf{V}_3 = \mathbf{0}$. Thus, this set is linearly dependent.

Example 2 Is the set $\{[1, \ 2], [\ 3, \ 4]\}$ linearly independent?

Solution The vector equation is

$$c_1[1 \ \ 2] + c_2[3 \ \ 4] = [0 \ \ 0].$$

This equation can be rewritten as

$$[c_1 \;\; 2c_1] + [3c_2 \;\; 4c_2] = [0 \;\; 0]$$

or as

$$[c_1 + 3c_2 \;\; 2c_1 + 4c_2] = [0 \;\; 0].$$

Equating components, we see that this vector equation is equivalent to the system

$$\begin{aligned} c_1 + 3c_2 &= 0, \\ 2c_1 + 4c_2 &= 0. \end{aligned}$$

Using Gaussian elimination, we find that the only solution to this system is $c_1 = c_2 = 0$, hence the original set of vectors is linearly independent. ■

Although we have worked exclusively with row vectors, the above definitions are equally applicable to column vectors.

Example 3 Is the set

$$\left\{ \begin{bmatrix} 2 \\ 6 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 8 \\ 16 \\ -3 \end{bmatrix} \right\}$$

linearly independent?

Solution Consider the vector equation

$$c_1 \begin{bmatrix} 2 \\ 6 \\ -2 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} + c_3 \begin{bmatrix} 8 \\ 16 \\ -3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{8}$$

This equation can be rewritten as

$$\begin{bmatrix} 2c_1 \\ 6c_1 \\ -2c_1 \end{bmatrix} + \begin{bmatrix} 3c_2 \\ c_2 \\ 2c_2 \end{bmatrix} + \begin{bmatrix} 8c_3 \\ 16c_3 \\ -3c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or as

$$\begin{bmatrix} 2c_1 + 3c_2 + 8c_3 \\ 6c_1 + c_2 + 16c_3 \\ -2c_1 + 2c_2 - 3c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

By equating components, we see that this vector equation is equivalent to the system

$$\begin{aligned} 2c_1 + 3c_2 + 8c_3 &= 0, \\ 6c_1 + c_2 + 16c_3 &= 0, \\ -2c_1 + 2c_2 - 3c_3 &= 0. \end{aligned}$$

By using Gaussian elimination, we find that the solution to this system is $c_1 = \left(-\frac{5}{2}\right) c_3$, $c_2 = -c_3$, c_3 arbitrary. Thus, choosing $c_3 = 2$, we obtain $c_1 = -5$, $c_2 = -2$, $c_3 = 2$ as a particular nonzero set of constants that satisfies (8); hence, the original vectors are linearly dependent. ■

Example 4 Is the set

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \end{bmatrix} \right\}$$

linearly independent?

Solution Consider the vector equation

$$c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 5 \\ 7 \end{bmatrix} + c_3 \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This is equivalent to the system

$$\begin{aligned} c_1 + 5c_2 - 3c_3 &= 0, \\ 2c_1 + 7c_2 + c_3 &= 0. \end{aligned}$$

By using Gaussian elimination, we find that the solution to this system is $c_1 = (-26/3)c_3$, $c_2 = (7/3)c_3$, c_3 arbitrary. Hence a particular nonzero solution is found by choosing $c_3 = 3$; then $c_1 = -26$, $c_2 = 7$, and, therefore, the vectors are linearly dependent. ■

We conclude this section with a few important theorems on linear independence and dependence.

Theorem 1 *A set of vectors is linearly dependent if and only if one of the vectors is a linear combination of the others.*

Proof. Let $\{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n\}$ be a linearly dependent set. Then there exist scalars $c_1, c_2, \ldots, c_n$, not all zero, such that (7) is satisfied. Assume $c_1 \neq 0$. (Since at least one of the c's must differ from zero, we lose no generality in assuming it is c_1). Equation (7) can be rewritten as

$$c_1\mathbf{V}_1 = -c_2\mathbf{V}_2 - c_3\mathbf{V}_3 - \cdots - c_n\mathbf{V}_n,$$

or as

$$\mathbf{V}_1 = -\frac{c_2}{c_1}\mathbf{V}_2 - \frac{c_3}{c_1}\mathbf{V}_3 - \cdots - \frac{c_n}{c_1}\mathbf{V}_n.$$

Thus, $\mathbf{V}_1$ is a linear combination of $\mathbf{V}_2, \mathbf{V}_3, \ldots, \mathbf{V}_n$. To complete the proof, we must show that if one vector is a linear combination of the others, then the set is linearly dependent. We leave this as an exercise for the student (see Problem 36.)

Observation 1 In order for a set of vectors to be linearly dependent, it is not necessary for *every* vector to be a linear combination of the others, only that there exists *one* vector that is a linear combination of the others. For example, consider the vectors $[1 \quad 0]$, $[2 \quad 0]$, $[0 \quad 1]$. Here, $[0, \quad 1]$ cannot be written as a linear combination of the other two vectors; however, $[2 \quad 0]$ can be written as a linear combination of $[1 \quad 0]$ and $[0 \quad 1]$, namely, $[2 \quad 0] = 2[1 \quad 0] + 0[0 \quad 1]]$; hence, the vectors are linearly dependent. □

Theorem 2 *The set consisting of the single vector* $\mathbf{V}_1$ *is a linearly independent set if and only if* $\mathbf{V}_1 \neq 0$.

Proof. Consider the equation $c_1\mathbf{V}_1 = \mathbf{0}$. If $\mathbf{V}_1 \neq \mathbf{0}$, then the only way this equation can be valid is if $c_1 = 0$; hence, the set is linearly independent. If $\mathbf{V}_1 = \mathbf{0}$, then any $c_1 \neq 0$ will satisfy the equation; hence, the set is linearly dependent. □

Theorem 3 *Any set of vectors that contains the zero vector is linearly dependent.*

Proof. Consider the set $\{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n, \mathbf{0}\}$. Pick $c_1 = c_2 = \cdots = c_n = 0$, $c_{n+1} = 5$ (any other number will do). Then this is a set of scalars, not all zero, such that

$$c_1\mathbf{V}_1 + c_2\mathbf{V}_2 + \cdots + c_n\mathbf{V}_n + c_{n+1}\mathbf{0} = \mathbf{0};$$

hence, the set of vectors is linearly dependent. □

Theorem 4 *If a set of vectors is linearly independent, any subset of these vectors is also linearly independent.*

Proof. See Problem 37. □

Theorem 5 *If a set of vectors is linearly dependent, then any larger set, containing this set, is also linearly dependent.*

Proof. See Problem 38. □

Problems 2.5

In Problems 1 through 19, determine whether or not the given set is linearly independent.

1. $\{[1 \quad 0], [0 \quad 1]\}$.

2. $\{[1 \quad 1], [1 \quad -1]\}$.

3. $\{[2 \quad -4], [-3 \quad 6]\}$.

4. $\{[1 \quad 3], [2 \quad -1], [1 \quad 1]\}$.

5. $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}\right\}$.

6. $\left\{\begin{bmatrix}1\\-1\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}\right\}$.

7. $\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$.

8. $\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\2\end{bmatrix}, \begin{bmatrix}2\\0\\1\end{bmatrix}\right\}$.

9. $\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\1\end{bmatrix}\right\}$.

10. $\left\{\begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}3\\2\\1\end{bmatrix}, \begin{bmatrix}2\\1\\3\end{bmatrix}\right\}$.

11. $\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}3\\2\\1\end{bmatrix}, \begin{bmatrix}2\\1\\3\end{bmatrix}\right\}$.

12. $\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}3\\2\\1\end{bmatrix}, \begin{bmatrix}2\\1\\3\end{bmatrix}, \begin{bmatrix}-1\\2\\3\end{bmatrix}\right\}$.

13. $\left\{\begin{bmatrix}4\\5\\1\end{bmatrix}, \begin{bmatrix}3\\0\\2\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$.

14. $\{[1 \quad 1 \quad 0], [1 \quad -1 \quad 0]\}$.

15. $\{[1 \quad 2 \quad 3], [-3 \quad -6 \quad -9]\}$.

16. $\{[10 \quad 20 \quad 20], [10 \quad -10 \quad 10], [10 \quad 20 \quad 10]\}$.

17. $\{[10 \quad 20 \quad 20], [10 \quad -10 \quad 10], [10 \quad 20 \quad 10], [20 \quad 10 \quad 20]\}$.

18. $\{[2 \quad 1 \quad 1], [3 \quad -1 \quad 4], [1 \quad 3 \quad -2]\}$.

19. $\left\{\begin{bmatrix}2\\1\\1\\3\end{bmatrix}, \begin{bmatrix}4\\-1\\2\\-1\end{bmatrix}, \begin{bmatrix}8\\1\\4\\5\end{bmatrix}\right\}$.

20. Express the vector

$$\begin{bmatrix}2\\1\\2\end{bmatrix}$$

as a linear combination of

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

21. Can the vector $[2 \quad 3]$ be expressed as a linear combination of the vectors given in (a) Problem 1, (b) Problem 2, or (c) Problem 3?

22. Can the vector $[1 \quad 1 \quad 1]^{\mathsf{T}}$ be expressed as a linear combination of the vectors given in (a) Problem 7, (b) Problem 8, or (c) Problem 9?

23. Can the vector $[2 \quad 0 \quad 3]^{\mathsf{T}}$ be expressed as a linear combination of the vectors given in Problem 8?

24. A set of vectors S is a *spanning set* for another set of vectors R if every vector in R can be expressed as a linear combination of the vectors in S. Show that the vectors given in Problem 1 are a spanning set for all two-dimensional row vectors. *Hint:* Show that for any arbitrary real numbers a and b, the vector $[a \quad b]$ can be expressed as a linear combination of the vectors in Problem 1.

25. Show that the vectors given in Problem 2 are a spanning set for all two-dimensional row vectors.

26. Show that the vectors given in Problem 3 are not a spanning set for all two-dimensional row vectors.

27. Show that the vectors given in Problem 3 are a spanning set for all vectors of the form $[a \quad -2a]$, where a designates any real number.

28. Show that the vectors given in Problem 4 are a spanning set for all two-dimensional row vectors.

29. Determine whether the vectors given in Problem 7 are a spanning set for all three-dimensional column vectors.

30. Determine whether the vectors given in Problem 8 are a spanning set for all three-dimensional column vectors.

31. Determine whether the vectors given in Problem 8 are a spanning set for vectors of the form $[a \quad 0 \quad a]^{\mathsf{T}}$, where a denotes an arbitrary real number.

32. A set of vectors S is a *basis* for another set of vectors R if S is a spanning set for R and S is linearly independent. Determine which, if any, of the sets given in Problems 1 through 4 are a basis for the set of all two dimensional row vectors.

33. Determine which, if any, of the sets given in Problems 7 through 12 are a basis for the set of all three dimensional column vectors.

34. Prove that the columns of the 3×3 identity matrix form a basis for the set of all three dimensional column vectors.

35. Prove that the rows of the 4×4 identity matrix form a basis for the set of all four dimensional row vectors.

36. Finish the proof of Theorem 1. (*Hint*: Assume that $\mathbf{V}_1$ can be written as a linear combination of the other vectors.)

37. Prove Theorem 4.

38. Prove Theorem 5.

39. Prove that the set of vectors $\{\mathbf{x}, k\mathbf{x}\}$ is linearly dependent for any choice of the scalar k.

40. Prove that if $\mathbf{x}$ and $\mathbf{y}$ are linearly independent, then so too are $\mathbf{x} + \mathbf{y}$ and $\mathbf{x} - \mathbf{y}$.

41. Prove that if the set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is linearly independent then so too is the set $\{k_1\mathbf{x}_1, k_2\mathbf{x}_2, \ldots, k_n\mathbf{x}_n\}$ for any choice of the *non-zero* scalars $k_1, k_2, \ldots, k_n$.

42. Let $\mathbf{A}$ be an $n \times n$ matrix and let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ and $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$ be two sets of n-dimensional column vectors having the property that $\mathbf{A}\mathbf{x}_i = \mathbf{y}_i = 1, 2, \ldots, k$. Show that the set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is linearly independent if the set $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$ is.

2.6 Rank

If we interpret each row of a matrix as a row vector, the elementary row operations are precisely the operations used to form linear combinations; namely, multiplying vectors (rows) by scalars and adding vectors (rows) to other vectors (rows). This observation allows us to develop a straightforward matrix procedure for determining when a set of vectors is linearly independent. It rests on the concept of rank.

Definition 1 The *row rank* of a matrix is the maximum number of linearly independent vectors that can be formed from the rows of that matrix, considering each row as a separate vector. Analogically, the *column rank* of a matrix is the maximum number of linearly independent columns, considering each column as a separate vector.

Row rank is particularly easy to determine for matrices in row-reduced form.

Theorem 1 *The row rank of a row-reduced matrix is the number of nonzero rows in that matrix.*

Proof. We must prove two facts: First, that the nonzero rows, considered as vectors, form a linearly independent set, and second, that every larger set is linearly dependent. Consider the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r = \mathbf{0}, \tag{9}$$

where $\mathbf{v}_1$ is the first nonzero row, $\mathbf{v}_2$ is the second nonzero row, ..., and $\mathbf{v}_r$ is the last nonzero row of a row-reduced matrix. The first nonzero element in the first nonzero row of a row-reduced matrix must be unity. Assume it appears in column j. Then, no other rows have a nonzero element in that column. Consequently, when the left side of Eq. (9) is computed, it will have c_1 as its jth component. Since the right side of Eq. (9) is the zero vector, it follows that $c_1 = 0$. A similar argument then shows iteratively that $c_2, \ldots, c_r$, are all zero. Thus, the nonzero rows are linearly independent.

If all the rows of the matrix are nonzero, then they must comprise a maximum number of linearly independent vectors, because the row rank cannot be greater than the number of rows in the matrix. If there are zero rows in the row-reduced matrix, then it follows from Theorem 3 of Section 2.5 that including them could not increase the number of linearly independent rows. Thus, the largest number of linearly independent rows comes from including just the nonzero rows. □

Example 1 Determine the row rank of the matrix

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

Solution $\mathbf{A}$ is in row-reduced form. Since it contains three nonzero rows, its row rank is three. ■

The following two theorems, which are proved in the Final Comments to this chapter, are fundamental.

Theorem 2 *The row rank and column rank of a matrix are equal.*

For any matrix $\mathbf{A}$, we call this common number the *rank* of $\mathbf{A}$ and denote it by $r(\mathbf{A})$.

Theorem 3 *If $\mathbf{B}$ is obtained from $\mathbf{A}$ by an elementary row (or column) operation, then $r(\mathbf{B}) = r(\mathbf{A})$.*

Theorems 1 through 3 suggest a useful procedure for determining the rank of any matrix: Simply use elementary row operations to transform the given matrix to row-reduced form, and then count the number of nonzero rows.

Example 2 Determine the rank of

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 \\ 2 & -1 & 1 \\ 3 & 2 & 5 \\ 5 & 15 & 20 \end{bmatrix}.$$

Solution In Example 2 of Section 2.3, we transferred this matrix into the row-reduced form

$$\begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

This matrix has two nonzero rows so its rank, as well as that of $\mathbf{A}$, is two. ■

Example 3 Determine the rank of

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 3 & -1 & -6 \\ 3 & -2 & -4 & -2 \end{bmatrix}.$$

Solution In Example 3 of Section 2.3, we transferred this matrix into the row-reduced form

$$\begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 12 \\ 0 & 0 & 1 & 5 \end{bmatrix}.$$

This matrix has three nonzero rows so its rank, as well as that of $\mathbf{B}$, is three. ■

A similar procedure can be used for determining whether a set of vectors is linearly independent: Form a matrix in which each row is one of the vectors in the given set, and then determine the rank of that matrix. If the rank equals the number of vectors, the set is linearly independent; if not, the set is linearly dependent. In either case, the rank is the maximal number of linearly independent vectors that can be formed from the given set.

Example 4 Determine whether the set

$$\left\{ \begin{bmatrix} 2 \\ 6 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 8 \\ 16 \\ -3 \end{bmatrix} \right\}$$

is linearly independent.

Solution We consider the matrix

$$\begin{bmatrix} 2 & 6 & -2 \\ 3 & 1 & 2 \\ 8 & 16 & -3 \end{bmatrix}.$$

Reducing this matrix to row-reduced form, we obtain

$$\begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & -\frac{5}{8} \\ 0 & 0 & 0 \end{bmatrix}.$$

This matrix has two nonzero rows, so its rank is two. Since this is less than the number of vectors in the given set, that set is linearly dependent.

We can say even more: The original set of vectors contains a subset of two linearly independent vectors, the same number as the rank. Also, since no row interchanges were involved in the transformation to row-reduced form, we can conclude that the third vector is linear combination of the first two. ■

Example 5 Determine whether the set

$$\{[0\ \ 1\ \ 2\ \ 3\ \ 0], [1\ \ 3\ \ -1\ \ 2\ \ 1],$$
$$[2\ \ 6\ \ -1\ \ -3\ \ 1], [4\ \ 0\ \ 1\ \ 0\ \ 2]\}$$

is linearly independent.

Solution We consider the matrix

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 0 \\ 1 & 3 & -1 & 2 & 1 \\ 2 & 6 & -1 & -3 & 1 \\ 4 & 0 & 1 & 0 & 2 \end{bmatrix},$$

which can be reduced (after the first two rows are interchanged) to the row-reduced form

$$\begin{bmatrix} 1 & 3 & -1 & 2 & 1 \\ 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & -7 & -1 \\ 0 & 0 & 0 & 1 & \frac{27}{175} \end{bmatrix}.$$

This matrix has four nonzero rows, hence its rank is four, which is equal to the number of vectors in the given set. Therefore, the set is linearly independent. ■

Example 6 Can the vector

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

be written as a linear combination of the vectors

$$\begin{bmatrix} 3 \\ 6 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 \\ 4 \end{bmatrix}?$$

Solution The matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 6 \\ 2 & 4 \end{bmatrix}$$

can be transformed into the row-reduced form

$$\begin{bmatrix} 3 & 6 \\ 0 & 0 \end{bmatrix},$$

which has rank one; hence **A** has just one linearly independent row vector. In contrast, the matrix

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 3 & 6 \\ 2 & 4 \end{bmatrix}$$

can be transformed into the row-reduced form,

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

which has rank two; hence **B** has two linearly independent row vectors. Since **B** is precisely **A** with one additional row, it follows that the additional row $[1, \quad 1]^{\mathsf{T}}$ is independent of the other two and, therefore, cannot be written as a linear combination of the other two vectors. ■

We did not have to transform **B** in Example 6 into row-reduced form to determine whether the three-vector set was linearly independent. There is a more direct approach. Since **B** has only two columns, its column rank must be less than or equal to two (why?). Thus, the column rank is less than three. It follows from Theorem 3 that the row rank of **B** is less than three, so the three vectors must be linearly dependent. Generalizing this reasoning, we deduce one of the more important results in linear algebra.

Theorem 4 *In an n-dimensional vector space, every set of $n + 1$ vectors is linearly dependent.*

Problems 2.6

In Problems 1–5, find the rank of the given matrix.

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

2. $\begin{bmatrix} 4 & 1 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

3. $\begin{bmatrix} 1 & 4 & -2 \\ 2 & 8 & -4 \\ -1 & -4 & 2 \end{bmatrix}$.

4. $\begin{bmatrix} 1 & 2 & 4 & 2 \\ 1 & 1 & 3 & 2 \\ 1 & 2 & 4 & 2 \end{bmatrix}$.

5. $\begin{bmatrix} 1 & 7 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

In Problems 6 through 22, use rank to determine whether the given set of vectors is linearly independent.

6. $\{[1 \quad 0], [0 \quad 1]\}$.

7. $\{[1 \quad 1], [1 \quad -1]\}$.

8. $\{[2 \quad -4], [-3 \quad 6]\}$.

9. $\{[1 \quad 3], [2 \quad -1], [1 \quad 1]\}$.

10. $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\}$.

11. $\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$.

12. $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$.

13. $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$.

14. $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$.

15. $\left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \right\}$.

16. $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \right\}$.

17. $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} \right\}$.

18. $\{[1 \quad 1 \quad 0], [1 \quad -1 \quad 0]\}$.

19. $\{[1 \quad 2 \quad 3], [-3 \quad -6 \quad -9]\}$.

20. $\{[10 \quad 20 \quad 20], [10 \quad -10 \quad 10], [10 \quad 20 \quad 10]\}$.

21. $\{[10 \quad 20 \quad 20], [10 \quad -10 \quad 10], [10 \quad 20 \quad 10], [20 \quad 10 \quad 20]\}$.

22. $\{[2 \quad 1 \quad 1], [3 \quad -1 \quad 4], [1 \quad 3 \quad -2]\}$.

23. Can the vector $[2 \quad 3]$ be expressed as a linear combination of the vectors given in (a) Problem 6, (b) Problem 7, or (c) Problem 8?

24. Can the vector $[1 \quad 1 \quad 1]^{\mathrm{T}}$ be expressed as a linear combination of the vectors given in (a) Problem 12, (b) Problem 13, or (c) Problem 14?

25. Can the vector $[2 \quad 0 \quad 3]^{\mathrm{T}}$ be expressed as a linear combination of the vectors given in Problem 13?

26. Can $[3 \quad 7]$ be written as a linear combination of the vectors $[1 \quad 2]$ and $[3 \quad 2]$?

27. Can $[3 \quad 7]$ be written as a linear combination of the vectors $[1 \quad 2]$ and $[4 \quad 8]$?

28. Find a maximal linearly independent subset of the vectors given in Problem 9.

29. Find a maximal linearly independent subset of the vectors given in Problem 13.

30. Find a maximal linearly independent subset of the set.

$$[1 \quad 2 \quad 4 \quad 0], [2 \quad 4 \quad 8 \quad 0], [1 \quad -1 \quad 0 \quad 1], [4 \quad 2 \quad 8 \quad 2], [4 \quad -1 \quad 4 \quad 3].$$

31. What is the rank of the zero matrix?

32. Show $r(\mathbf{A}^{\mathrm{T}}) = r(\mathbf{A})$.

2.7 Theory of Solutions

Consider once again the system $\mathbf{Ax} = \mathbf{b}$ of m equations and n unknowns given in Eq. (2). Designate the n columns of $\mathbf{A}$ by the vectors $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n$. Then Eq. (2) can be rewritten in the vector form

$$x_1\mathbf{V}_1 + x_2\mathbf{V}_2 + \cdots + x_n\mathbf{V}_n = \mathbf{b}. \tag{10}$$

Example 1 Rewrite the following system in the vector form (10):

$$x - 2y + 3z = 7,$$
$$4x + 5y - 6z = 8.$$

Solution

$$x\begin{bmatrix} 1 \\ 4 \end{bmatrix} + y\begin{bmatrix} -2 \\ 5 \end{bmatrix} + z\begin{bmatrix} 3 \\ -6 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix} \quad \blacksquare$$

Thus, finding solutions to (1) and (2) is equivalent to finding scalars $x_1, x_2, \ldots, x_n$ that satisfy (10). This, however, is asking precisely the question "Is the vector $\mathbf{b}$ a linear combination of $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n$?" If $\mathbf{b}$ is a linear combination of $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n$, then there will exist scalars $x_1, x_2, \ldots, x_n$ that satisfy (10) and the system is consistent. If $\mathbf{b}$ is not a linear combination of these vectors, that is, if $\mathbf{b}$ is linearly independent of the vectors $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n$, then no scalars $x_1, x_2, \ldots, x_n$ will exist that satisfy (10) and the system is inconsistent.

Taking a hint from Example 6 of Section 2.6, we have the following theorem.

Theorem 1 *The system* $\mathbf{Ax} = \mathbf{b}$ *is consistent if and only if* $r(\mathbf{A}) = r(\mathbf{A^b})$.

Once a system is deemed consistent, the following theorem specifies the number of solutions.

Theorem 2 *If the system* $\mathbf{Ax} = \mathbf{b}$ *is consistent and* $r(\mathbf{A}) = k$ *then the solutions are expressible in terms of* $n - k$ *arbitrary unknowns (where* n *represents the number of unknowns in the system).*

Theorem 2 is almost obvious. To determine the rank of $\mathbf{A^b}$, we must reduce it to row-reduced form. The rank is the number of nonzero rows. With Gaussian elimination, we use each nonzero row to solve for the variable associated with the first nonzero entry in it. Thus, each nonzero row defines one variable, and all other variables remain arbitrary.

Example 2 Discuss the solutions of the system

$$\begin{aligned} x + y - z &= 1, \\ x + y - z &= 0. \end{aligned}$$

Solution

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{A^b} = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix}.$$

Here, $r(\mathbf{A}) = 1$, $r(\mathbf{A^b}) = 2$. Thus, $r(\mathbf{A}) \neq r(\mathbf{A^b})$ and no solution exists. ■

Example 3 Discuss the solutions of the system

$$\begin{aligned} x + y + w &= 3, \\ 2x + 2y + 2w &= 6, \\ -x - y - w &= -3. \end{aligned}$$

Solution

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 6 \\ -3 \end{bmatrix}, \quad \mathbf{A^b} = \begin{bmatrix} 1 & 1 & 1 & 3 \\ 2 & 2 & 2 & 6 \\ -1 & -1 & -1 & -3 \end{bmatrix}.$$

Here $r(\mathbf{A}) = r(\mathbf{A^b}) = 1$; hence, the system is consistent. In this case, $n = 3$ and $k = 1$; thus, the solutions are expressible in terms of $3 - 1 = 2$ arbitrary unknowns. Using Gaussian elimination, we find that the solution is $x = 3 - y - w$ where y and w are both arbitrary. ■

Example 4 Discuss the solutions of the system

$$\begin{aligned} 2x - 3y + z &= -1, \\ x - y + 2z &= 2, \\ 2x + y - 3z &= 3. \end{aligned}$$

Solution

$$\mathbf{A} = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -1 & 2 \\ 2 & 1 & -3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{A^b} = \begin{bmatrix} 2 & -3 & 1 & -1 \\ 1 & -1 & 2 & 2 \\ 2 & 1 & -3 & 3 \end{bmatrix}.$$

Here $r(\mathbf{A}) = r(\mathbf{A^b}) = 3$, hence the system is consistent. Since $n = 3$ and $k = 3$, the solution will be in $n - k = 0$ arbitrary unknowns. Thus, the solution is unique (none of the unknowns are arbitrary) and can be obtained by Gaussian elimination as $x = y = 2, z = 1$. ■

Example 5 Discuss the solutions of the system

$$\begin{aligned} x + y - 2z &= 1, \\ 2x + y + z &= 2, \\ 3x + 2y - z &= 3, \\ 4x + 2y + 2z &= 4. \end{aligned}$$

Solution

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

Here $r(\mathbf{A}) = r(\mathbf{A^b}) = 2$. Thus, the system is consistent and the solutions will be in terms of $3 - 2 = 1$ arbitrary unknowns. Using Gaussian elimination, we find that the solution is $x = 1 - 3z$, $y = 5z$, and z is arbitrary. ■

In a consistent system, the solution is unique if $k = n$. If $k \neq n$, the solution will be in terms of arbitrary unknowns. Since these arbitrary unknowns can be chosen to be any constants whatsoever, it follows that there will be an infinite number of solutions. Thus, a consistent system will possess exactly one solution or an infinite number of solutions; there is no inbetween.

A homogeneous system of simultaneous linear equations has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0, \end{aligned} \tag{11}$$

or the matrix form

$$\mathbf{Ax} = 0. \tag{12}$$

Since Eq. (12) is a special case of Eq. (2) with $\mathbf{b} = \mathbf{0}$, all the theory developed for the system $\mathbf{Ax} = \mathbf{b}$ remains valid. Because of the simplified structure of a homogeneous system, one can draw conclusions about it that are not valid for a nonhomogeneous system. For instance, a homogeneous system is always consistent. To verify this statement, note that $x_1 = x_2 = \cdots = x_n = 0$ is always a solution to Eq. (12). Such a solution is called the *trivial solution*. It is, in general, the *nontrivial solutions* (solutions in which one or more of the unknowns is different from zero) that are of the greatest interest.

It follows from Theorem 2, that if the rank of $\mathbf{A}$ is less than n(n being the number of unknowns), then the solution will be in terms of arbitrary unknowns. Since these arbitrary unknowns can be assigned nonzero values, it follows that nontrivial solutions exist. On the other hand, if the rank of $\mathbf{A}$ equals n, then the solution will be unique, and, hence, must be the trivial solution (why?). Thus, it follows that:

Theorem 3 *The homogeneous system (12) will admit nontrivial solutions if and only if* $r(\mathbf{A}) \neq n$.

Problems 2.7

In Problems 1–9, discuss the solutions of the given system in terms of consistency and number of solutions. Check your answers by solving the systems wherever possible.

1. $\begin{aligned} x - 2y &= 0, \\ x + y &= 1, \\ 2x - y &= 1. \end{aligned}$

2. $\begin{aligned} x + y &= 0, \\ 2x - 2y &= 1, \\ x - y &= 0. \end{aligned}$

3. $\begin{aligned} x + y + z &= 1, \\ x - y + z &= 2, \\ 3x + y + 3z &= 4. \end{aligned}$

4. $\begin{aligned} x + 3y + 2z - w &= 2, \\ 2x - y + z + w &= 3. \end{aligned}$

5. $$\begin{aligned} 2x - y + z &= 0, \\ x + 2y - z &= 4, \\ x + y + z &= 1. \end{aligned}$$

6. $$\begin{aligned} 2x + 3y &= 0, \\ x - 4y &= 0, \end{aligned}$$

7. $$\begin{aligned} x - y + 2z &= 0, \\ 2x + 3y - z &= 0, \\ -2x + 7y - 7z &= 0. \end{aligned}$$

8. $$\begin{aligned} x - y + 2z &= 0, \\ 2x - 3y - z &= 0, \\ -2x + 7y - 9z &= 0. \end{aligned}$$

9. $$\begin{aligned} x - 2y + 3z + 3w &= 0, \\ y - 2z + 2w &= 0, \\ x + y - 3z + 9w &= 0. \end{aligned}$$

2.8 Final Comments on Chapter 2

We would like to show that the column rank of a matrix equals its row rank, and that an elementary row operation of any kind does not alter the rank.

Lemma 1 *If* $\mathbf{B}$ *is obtained from* $\mathbf{A}$ *by interchanging two columns of* $\mathbf{A}$*, then both* $\mathbf{A}$ *and* $\mathbf{B}$ *have the same column rank.*

Proof. The set of vectors formed from the columns of $\mathbf{A}$ is identical to the set formed from the columns of $\mathbf{B}$, and, therefore, the two matrices must have the same column rank. □

Lemma 2 *If* $\mathbf{Ax} = \mathbf{0}$ *and* $\mathbf{Bx} = \mathbf{0}$ *have the same set of solutions, then the column rank of* $\mathbf{A}$ *is less than or equal to the column rank of* $\mathbf{B}$.

Proof. Let the order of $\mathbf{A}$ be $m \times n$. Then, the system $\mathbf{Ax} = \mathbf{0}$ is a set of m equations in the n unknowns $x_1, x_2, \ldots, x_n$, which has the vector form

$$x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_n\mathbf{A}_n = \mathbf{0}, \tag{13}$$

where $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ denote the columns of $\mathbf{A}$. Similarly, the system $\mathbf{Bx} = \mathbf{0}$ has the vector form

$$x_1\mathbf{B}_1 + x_2\mathbf{B}_2 + \cdots + x_n\mathbf{B}_n = \mathbf{0}. \tag{14}$$

We shall assume that the column rank of $\mathbf{A}$ is greater than the column rank of $\mathbf{B}$ and show that this assumption leads to a contradiction. It will then follow that the reverse must be true, which is precisely what we want to prove.

Denote the column rank of $\mathbf{A}$ as a and the column rank of $\mathbf{B}$ as b. We assume that $a > b$. Since the column rank of $\mathbf{A}$ is a, there must exist a columns of $\mathbf{A}$ that are linearly independent. If these columns are not the first a columns, rearrange the order of the columns so they are. Lemma 1 guarantees such reorderings do not alter the column rank. Thus, $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_a$ are linearly independent. Since a is assumed greater than b, we know that the first a columns of $\mathbf{B}$ are not linearly independent. Since they are linearly dependent, there must exist constants $c_1, c_2, \ldots, c_a$ — not all zero — such that

$$c_1\mathbf{B}_1 + c_2\mathbf{B}_2 + \cdots + c_a\mathbf{B}_a = \mathbf{0}.$$

It then follows that

$$c_1\mathbf{B}_1 + c_2\mathbf{B}_2 + \cdots + c_a\mathbf{B}_a + 0\mathbf{B}_{a+1} + \cdots + 0\mathbf{B}_n = \mathbf{0},$$

from which we conclude that

$$x_1 = c_1, \quad x_2 = c_2, \quad \ldots, x_a = c_a, \quad x_{a+1} = 0, \quad \ldots, x_n = 0.$$

is a solution of Eq. (14). Since every solution to Eq. (14) is also a solution to Eq. (12), we have

$$c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \cdots + c_a\mathbf{A}_a + 0\mathbf{A}_{a+1} + \cdots + 0\mathbf{A}_n = \mathbf{0},$$

or more simply

$$c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \cdots + c_a\mathbf{A}_a = \mathbf{0},$$

where all the c's are not all zero. But this implies that the first a columns of $\mathbf{A}$ are linearly dependent, which is a contradiction of the assumption that they were linearly independent. □

Lemma 3 *If* $\mathbf{Ax} = \mathbf{0}$ *and* $\mathbf{Bx} = \mathbf{0}$ *have the same set of solutions, then* $\mathbf{A}$ *and* $\mathbf{B}$ *have the same column rank.*

Proof. If follows from Lemma 2 that the column rank of $\mathbf{A}$ is less than or equal to the column rank of $\mathbf{B}$. By reversing the roles of $\mathbf{A}$ and $\mathbf{B}$, we can also conclude from Lemma 2 that the column rank of $\mathbf{B}$ is less than or equal to the column rank of $\mathbf{A}$. As a result, the two column ranks must be equal. □

Theorem 1 *An elementary row operation does not alter the column rank of a matrix.*

Proof. Denote the original matrix as $\mathbf{A}$, and let $\mathbf{B}$ denote a matrix obtained by applying an elementary row operation to $\mathbf{A}$; and consider the two homogeneous systems $\mathbf{Ax} = \mathbf{0}$ and $\mathbf{Bx} = \mathbf{0}$. Since elementary row operations do not alter solutions, both of these systems have the same solution set. Theorem 1 follows immediately from Lemma 3. □

Lemma 4 *The column rank of a matrix is less than or equal to its row rank.*

Proof. Denote rows of $\mathbf{A}$ by $\mathbf{A}_1, \mathbf{A}_2, \ldots \mathbf{A}_m$, the column rank of matrix $\mathbf{A}$ by c and its row rank by r. There must exist r rows of $\mathbf{A}$ which are linearly independent. If these rows are not the first r rows, rearrange the order of the rows so they are. Theorem 1 guarantees such reorderings do not alter the column rank, and they certainly do not alter the row rank. Thus, $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_r$ are linearly independent. Define partitioned matrices $\mathbf{R}$ and $\mathbf{S}$ by

$$\mathbf{R} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_r \end{bmatrix} \quad \text{and} \quad \mathbf{S} = \begin{bmatrix} \mathbf{A}_{r+1} \\ \mathbf{A}_{r+2} \\ \vdots \\ \mathbf{A}_n \end{bmatrix}.$$

Then $\mathbf{A}$ has the partitioned form

$$\mathbf{A} = \begin{bmatrix} \mathbf{R} \\ \mathbf{S} \end{bmatrix}.$$

Every row of $\mathbf{S}$ is a linear combination of the rows of $\mathbf{R}$. Therefore, there exist constants t_{ij} such that

$$\begin{aligned} \mathbf{A}_{r+1} &= t_{r+1,1}\mathbf{A}_1 + t_{r+1,2}\mathbf{A}_2 + \cdots + t_{r+1,r}\mathbf{A}_r, \\ \mathbf{A}_{r+2} &= t_{r+2,1}\mathbf{A}_1 + t_{r+2,2}\mathbf{A}_2 + \cdots + t_{r+2,r}\mathbf{A}_r, \\ &\vdots \\ \mathbf{A}_n &= t_{n,1}\mathbf{A}_1 + t_{n,2}\mathbf{A}_2 + \cdots + t_{n,r}\mathbf{A}_r, \end{aligned}$$

which may be written in the matrix form

$$\mathbf{S} = \mathbf{TR},$$

where

$$\mathbf{T} = \begin{bmatrix} t_{r+1,1} & t_{r+1,2} & \cdots & t_{r+1,n} \\ t_{r+2,1} & t_{r+2,2} & \cdots & t_{r+2,n} \\ \vdots & \vdots & \vdots & \vdots \\ t_{n,1} & t_{n,2} & \cdots & t_{n,n} \end{bmatrix}.$$

Then, for any n-dimensional vector $\mathbf{x}$, we have

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{R} \\ \mathbf{S} \end{bmatrix}\mathbf{x} = \begin{bmatrix} \mathbf{Rx} \\ \mathbf{Sx} \end{bmatrix} = \begin{bmatrix} \mathbf{Rx} \\ \mathbf{TRx} \end{bmatrix}.$$

Thus, $\mathbf{Ax} = \mathbf{0}$ if and only if $\mathbf{Rx} = \mathbf{0}$. It follows from Lemma 3 that both $\mathbf{A}$ and $\mathbf{R}$ have the same column rank. But the columns of $\mathbf{R}$ are r-dimensional vectors, so its column rank cannot be larger than r. Thus,

$$c = \text{column rank of } \mathbf{A} = \text{column rank of } \mathbf{R} \leq r = \text{row rank of } \mathbf{A} \qquad \square$$

Lemma 5 *The row rank of a matrix is less than or equal to its column rank.*

Proof. By applying Lemma 4 to $\mathbf{A}^\mathsf{T}$, we conclude that the column rank of $\mathbf{A}^\mathsf{T}$ is less than or equal to the row rank of $\mathbf{A}^\mathsf{T}$. But since the columns of $\mathbf{A}^\mathsf{T}$ are the rows of $\mathbf{A}$ and vice-versa, the result follows immediately. $\square$

Theorem 2 *The row rank of a matrix equals its column rank.*

Proof. The result is immediate from Lemmas 4 and 5. $\square$

Theorem 3 *An elementary row operation does not alter the row rank of a matrix.*

Proof. This theorem is an immediate consequence of both Theorems 1 and 2. $\square$

This page intentionally left blank