CHAPTER 3

Systems of Linear Equations

3.1 Introduction

Systems of linear equations play an important and motivating role in the subject of linear algebra. In fact, many problems in linear algebra reduce to finding the solution of a system of linear equations. Thus, the techniques introduced in this chapter will be applicable to abstract ideas introduced later. On the other hand, some of the abstract results will give us new insights into the structure and properties of systems of linear equations.

All our systems of linear equations involve scalars as both coefficients and constants, and such scalars may come from any number field K. There is almost no loss in generality if the reader assumes that all our scalars are real numbers—that is, that they come from the real field $\mathbf{R}$.

3.2 Basic Definitions, Solutions

This section gives basic definitions connected with the solutions of systems of linear equations. The actual algorithms for finding such solutions will be treated later.

Linear Equation and Solutions

A *linear equation* in unknowns $x_1, x_2, \ldots, x_n$ is an equation that can be put in the *standard form*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{3.1}$$

where $a_1, a_2, \ldots, a_n$, and b are constants. The constant a_k is called the *coefficient* of x_k, and b is called the *constant term* of the equation.

A solution of the linear equation (3.1) is a list of values for the unknowns or, equivalently, a vector u in K^n, say

$$x_1 = k_1, \quad x_2 = k_2, \quad \ldots, \quad x_n = k_n \qquad \text{or} \qquad u = (k_1, k_2, \ldots, k_n)$$

such that the following statement (obtained by substituting k_i for x_i in the equation) is true:

$$a_1k_1 + a_2k_2 + \cdots + a_nk_n = b$$

In such a case we say that u *satisfies* the equation.

Remark: Equation (3.1) implicitly assumes there is an ordering of the unknowns. In order to avoid subscripts, we will usually use x, y for two unknowns; x, y, z for three unknowns; and x, y, z, t for four unknowns; they will be ordered as shown.

EXAMPLE 3.1 Consider the following linear equation in three unknowns x, y, z:

$$x + 2y - 3z = 6$$

We note that $x = 5, y = 2, z = 1$, or, equivalently, the vector $u = (5, 2, 1)$ is a solution of the equation. That is,

$$5 + 2(2) - 3(1) = 6 \qquad \text{or} \qquad 5 + 4 - 3 = 6 \qquad \text{or} \qquad 6 = 6$$

On the other hand, $w = (1, 2, 3)$ is not a solution, because on substitution, we do not get a true statement:

$$1 + 2(2) - 3(3) = 6 \qquad \text{or} \qquad 1 + 4 - 9 = 6 \qquad \text{or} \qquad -4 = 6$$

System of Linear Equations

A system of linear equations is a list of linear equations with the same unknowns. In particular, a system of m linear equations $L_1, L_2, \ldots, L_m$ in n unknowns $x_1, x_2, \ldots, x_n$ can be put in the *standard form*

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \tag{3.2}$$

where the a_{ij} and b_i are constants. The number a_{ij} is the *coefficient* of the unknown x_j in the equation L_i, and the number b_i is the *constant* of the equation L_i.

The system (3.2) is called an $m \times n$ (read: m by n) system. It is called a *square system* if $m = n$—that is, if the number m of equations is equal to the number n of unknowns.

The system (3.2) is said to be *homogeneous* if all the constant terms are zero—that is, if $b_1 = 0$, $b_2 = 0, \ldots, b_m = 0$. Otherwise the system is said to be *nonhomogeneous*.

A *solution* (or a *particular solution*) of the system (3.2) is a list of values for the unknowns or, equivalently, a vector u in K^n, which is a solution of each of the equations in the system. The set of all solutions of the system is called the *solution set* or the *general solution* of the system.

EXAMPLE 3.2 Consider the following system of linear equations:

$$\begin{aligned} x_1 + x_2 + 4x_3 + 3x_4 &= 5 \\ 2x_1 + 3x_2 + x_3 - 2x_4 &= 1 \\ x_1 + 2x_2 - 5x_3 + 4x_4 &= 3 \end{aligned}$$

It is a 3×4 system because it has three equations in four unknowns. Determine whether (a) $u = (-8, 6, 1, 1)$ and (b) $v = (-10, 5, 1, 2)$ are solutions of the system.

(a) Substitute the values of u in each equation, obtaining

$$\begin{array}{rcrcr} -8 + 6 + 4(1) + 3(1) = 5 & \text{or} & -8 + 6 + 4 + 3 = 5 & \text{or} & 5 = 5 \\ 2(-8) + 3(6) + 1 - 2(1) = 1 & \text{or} & -16 + 18 + 1 - 2 = 1 & \text{or} & 1 = 1 \\ -8 + 2(6) - 5(1) + 4(1) = 3 & \text{or} & -8 + 12 - 5 + 4 = 3 & \text{or} & 3 = 3 \end{array}$$

Yes, u is a solution of the system because it is a solution of each equation.

(b) Substitute the values of v into each successive equation, obtaining

$$\begin{array}{rcrcr} -10 + 5 + 4(1) + 3(2) = 5 & \text{or} & -10 + 5 + 4 + 6 = 5 & \text{or} & 5 = 5 \\ 2(-10) + 3(5) + 1 - 2(2) = 1 & \text{or} & -20 + 15 + 1 - 4 = 1 & \text{or} & -8 = 1 \end{array}$$

No, v is not a solution of the system, because it is not a solution of the second equation. (We do not need to substitute v into the third equation.)

The system (3.2) of linear equations is said to be *consistent* if it has one or more solutions, and it is said to be *inconsistent* if it has no solution. If the field K of scalars is infinite, such as when K is the real field $\mathbf{R}$ or the complex field $\mathbf{C}$, then we have the following important result.

THEOREM 3.1: Suppose the field K is infinite. Then any system $\mathscr{L}$ of linear equations has (i) a unique solution, (ii) no solution, or (iii) an infinite number of solutions.

This situation is pictured in Fig. 3-1. The three cases have a geometrical description when the system $\mathscr{L}$ consists of two equations in two unknowns (Section 3.4).

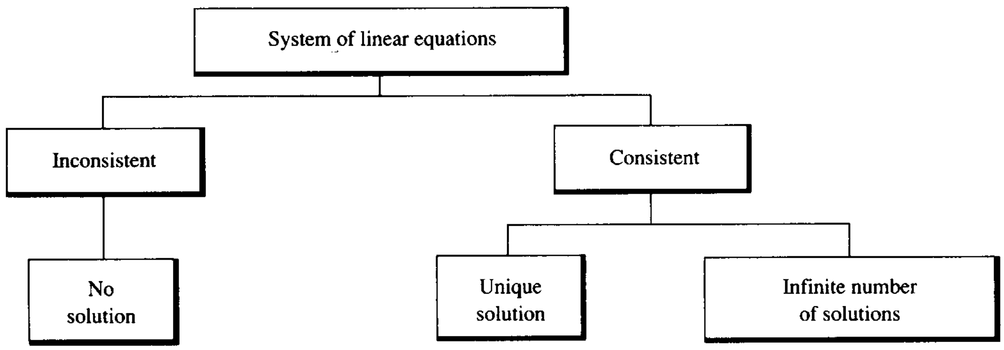

Figure 3-1

Augmented and Coefficient Matrices of a System

Consider again the general system (3.2) of m equations in n unknowns. Such a system has associated with it the following two matrices:

$$M = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

The first matrix M is called the *augmented matrix* of the system, and the second matrix A is called the *coefficient matrix*.

The coefficient matrix A is simply the matrix of coefficients, which is the augmented matrix M without the last column of constants. Some texts write $M = [A, B]$ to emphasize the two parts of M, where B denotes the column vector of constants. The augmented matrix M and the coefficient matrix A of the system in Example 3.2 are as follows:

$$M = \begin{bmatrix} 1 & 1 & 4 & 3 & 5 \\ 2 & 3 & 1 & -2 & 1 \\ 1 & 2 & -5 & 4 & 3 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 & 4 & 3 \\ 2 & 3 & 1 & -2 \\ 1 & 2 & -5 & 4 \end{bmatrix}$$

As expected, A consists of all the columns of M except the last, which is the column of constants.

Clearly, a system of linear equations is completely determined by its augmented matrix M, and vice versa. Specifically, each row of M corresponds to an equation of the system, and each column of M corresponds to the coefficients of an unknown, except for the last column, which corresponds to the constants of the system.

Degenerate Linear Equations

A linear equation is said to be *degenerate* if all the coefficients are zero—that is, if it has the form

$$0x_1 + 0x_2 + \dots + 0x_n = b \tag{3.3}$$

The solution of such an equation depends only on the value of the constant b. Specifically,

(i) If $b \neq 0$, then the equation has no solution.
(ii) If $b = 0$, then every vector $u = (k_1, k_2, \ldots, k_n)$ in K^n is a solution.

The following theorem applies.

THEOREM 3.2: Let $\mathscr{L}$ be a system of linear equations that contains a degenerate equation L, say with constant b.

(i) If $b \neq 0$, then the system $\mathscr{L}$ has no solution.

(ii) If $b = 0$, then L may be deleted from the system without changing the solution set of the system.

Part (i) comes from the fact that the degenerate equation has no solution, so the system has no solution. Part (ii) comes from the fact that every element in K^n is a solution of the degenerate equation.

Leading Unknown in a Nondegenerate Linear Equation

Now let L be a nondegenerate linear equation. This means one or more of the coefficients of L are not zero. By the *leading unknown* of L, we mean the first unknown in L with a nonzero coefficient. For example, x_3 and y are the leading unknowns, respectively, in the equations

$$0x_1 + 0x_2 + 5x_3 + 6x_4 + 0x_5 + 8x_6 = 7 \qquad \text{and} \qquad 0x + 2y - 4z = 5$$

We frequently omit terms with zero coefficients, so the above equations would be written as

$$5x_3 + 6x_4 + 8x_6 = 7 \qquad \text{and} \qquad 2y - 4z = 5$$

In such a case, the leading unknown appears first.

3.3 Equivalent Systems, Elementary Operations

Consider the system (3.2) of m linear equations in n unknowns. Let L be the linear equation obtained by multiplying the m equations by constants $c_1, c_2, \ldots, c_m$, respectively, and then adding the resulting equations. Specifically, let L be the following linear equation:

$$(c_1a_{11} + \cdots + c_ma_{m1})x_1 + \cdots + (c_1a_{1n} + \cdots + c_ma_{mn})x_n = c_1b_1 + \cdots + c_mb_m$$

Then L is called a *linear combination* of the equations in the system. One can easily show (Problem 3.43) that any solution of the system (3.2) is also a solution of the linear combination L.

EXAMPLE 3.3 Let L_1, L_2, L_3 denote, respectively, the three equations in Example 3.2. Let L be the equation obtained by multiplying L_1, L_2, L_3 by $3, -2, 4$, respectively, and then adding. Namely,

$$\begin{array}{rl}
3L_1: & 3x_1 + 3x_2 + 12x_3 + 9x_4 = 15 \\
-2L_2: & -4x_1 - 6x_2 - 2x_3 + 4x_4 = -2 \\
4L_1: & 4x_1 + 8x_2 - 20x_3 + 16x_4 = 12 \\
\hline
\text{(Sum) } L: & 3x_1 + 5x_2 - 10x_3 + 29x_4 = 25
\end{array}$$

Then L is a linear combination of L_1, L_2, L_3. As expected, the solution $u = (-8, 6, 1, 1)$ of the system is also a solution of L. That is, substituting u in L, we obtain a true statement:

$$3(-8) + 5(6) - 10(1) + 29(1) = 25 \qquad \text{or} \qquad -24 + 30 - 10 + 29 = 25 \qquad \text{or} \qquad 9 = 9$$

The following theorem holds.

THEOREM 3.3: Two systems of linear equations have the same solutions if and only if each equation in each system is a linear combination of the equations in the other system.

Two systems of linear equations are said to be *equivalent* if they have the same solutions. The next subsection shows one way to obtain equivalent systems of linear equations.

Elementary Operations

The following operations on a system of linear equations $L_1, L_2, \ldots, L_m$ are called *elementary operations*.

[E_1] Interchange two of the equations. We indicate that the equations L_i and L_j are interchanged by writing:

"Interchange L_i and L_j" or "$L_i \longleftrightarrow L_j$"

[E_2] Replace an equation by a nonzero multiple of itself. We indicate that equation L_i is replaced by kL_i (where $k \neq 0$) by writing

"Replace L_i by kL_i" or "$kL_i \to L_i$"

[E_3] Replace an equation by the sum of a multiple of another equation and itself. We indicate that equation L_j is replaced by the sum of kL_i and L_j by writing

"Replace L_j by $kL_i + L_j$" or "$kL_i + L_j \to L_j$"

The arrow $\to$ in [E_2] and [E_3] may be read as "replaces."

The main property of the above elementary operations is contained in the following theorem (proved in Problem 3.45).

THEOREM 3.4: Suppose a system of $\mathscr{M}$ of linear equations is obtained from a system $\mathscr{L}$ of linear equations by a finite sequence of elementary operations. Then $\mathscr{M}$ and $\mathscr{L}$ have the same solutions.

Remark: Sometimes (say to avoid fractions when all the given scalars are integers) we may apply [E_2] and [E_3] in one step; that is, we may apply the following operation:

[E] Replace equation L_j by the sum of kL_i and $k'L_j$ (where $k' \neq 0$), written

"Replace L_j by $kL_i + k'L_j$" or "$kL_i + k'L_j \to L_j$"

We emphasize that in operations [E_3] and [E], only equation L_j is changed.

Gaussian elimination, our main method for finding the solution of a given system of linear equations, consists of using the above operations to transform a given system into an equivalent system whose solution can be easily obtained.

The details of Gaussian elimination are discussed in subsequent sections.

3.4 Small Square Systems of Linear Equations

This section considers the special case of one equation in one unknown, and two equations in two unknowns. These simple systems are treated separately because their solution sets can be described geometrically, and their properties motivate the general case.

Linear Equation in One Unknown

The following simple basic result is proved in Problem 3.5.

THEOREM 3.5: Consider the linear equation $ax = b$.

(i) If $a \neq 0$, then $x = b/a$ is a unique solution of $ax = b$.

(ii) If $a = 0$, but $b \neq 0$, then $ax = b$ has no solution.

(iii) If $a = 0$ and $b = 0$, then every scalar k is a solution of $ax = b$.

EXAMPLE 3.4 Solve (a) $4x - 1 = x + 6$, (b) $2x - 5 - x = x + 3$, (c) $4 + x - 3 = 2x + 1 - x$.

(a) Rewrite the equation in standard form obtaining $3x = 7$. Then $x = \frac{7}{3}$ is the unique solution [Theorem 3.5(i)].

(b) Rewrite the equation in standard form, obtaining $0x = 8$. The equation has no solution [Theorem 3.5(ii)].

(c) Rewrite the equation in standard form, obtaining $0x = 0$. Then every scalar k is a solution [Theorem 3.5(iii)].

System of Two Linear Equations in Two Unknowns (2×2 System)

Consider a system of two nondegenerate linear equations in two unknowns x and y, which can be put in the standard form

$$\begin{aligned} A_1x + B_1y &= C_1 \\ A_2x + B_2y &= C_2 \end{aligned} \tag{3.4}$$

Because the equations are nondegenerate, A_1 and B_1 are not both zero, and A_2 and B_2 are not both zero.

The general solution of the system (3.4) belongs to one of three types as indicated in Fig. 3-1. If **R** is the field of scalars, then the graph of each equation is a line in the plane $\mathbf{R}^2$ and the three types may be described geometrically as pictured in Fig. 3-2. Specifically,

(1) *The system has exactly one solution.*
Here the two lines intersect in one point [Fig. 3-2(a)]. This occurs when the lines have distinct slopes or, equivalently, when the coefficients of x and y are not proportional:

$$\frac{A_1}{A_2} \neq \frac{B_1}{B_2} \qquad \text{or, equivalently,} \qquad A_1B_2 - A_2B_1 \neq 0$$

For example, in Fig. 3-2(a), $1/3 \neq -1/2$.

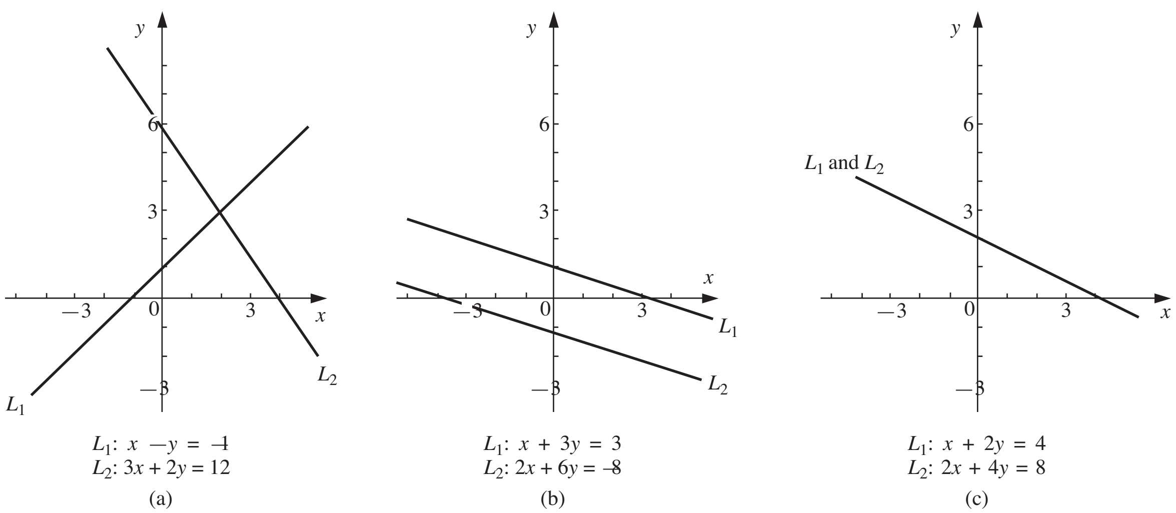

Figure 3-2

(2) *The system has no solution.*
Here the two lines are parallel [Fig. 3-2(b)]. This occurs when the lines have the same slopes but different y intercepts, or when

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} \neq \frac{C_1}{C_2}$$

For example, in Fig. 3-2(b), $1/2 = 3/6 \neq -3/8$.

(3) *The system has an infinite number of solutions.*
Here the two lines coincide [Fig. 3-2(c)]. This occurs when the lines have the same slopes and same y intercepts, or when the coefficients and constants are proportional,

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}$$

For example, in Fig. 3-2(c), $1/2 = 2/4 = 4/8$.

Remark: The following expression and its value is called a *determinant of order two*:

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} = A_1 B_2 - A_2 B_1$$

Determinants will be studied in Chapter 8. Thus, the system (3.4) has a unique solution if and only if the determinant of its coefficients is not zero. (We show later that this statement is true for any square system of linear equations.)

Elimination Algorithm

The solution to system (3.4) can be obtained by the process of elimination, whereby we reduce the system to a single equation in only one unknown. Assuming the system has a unique solution, this elimination algorithm has two parts.

ALGORITHM 3.1: The input consists of two nondegenerate linear equations L_1 and L_2 in two unknowns with a unique solution.

Part A. (Forward Elimination) Multiply each equation by a constant so that the resulting coefficients of one unknown are negatives of each other, and then add the two equations to obtain a new equation L that has only one unknown.

Part B. (Back-Substitution) Solve for the unknown in the new equation L (which contains only one unknown), substitute this value of the unknown into one of the original equations, and then solve to obtain the value of the other unknown.

Part A of Algorithm 3.1 can be applied to any system even if the system does not have a unique solution. In such a case, the new equation L will be degenerate and Part B will not apply.

EXAMPLE 3.5 (Unique Case). Solve the system

$$L_1:\ 2x - 3y = -8$$
$$L_2:\ 3x + 4y = \quad 5$$

The unknown x is eliminated from the equations by forming the new equation $L = -3L_1 + 2L_2$. That is, we multiply L_1 by -3 and L_2 by 2 and add the resulting equations as follows:

$$\begin{array}{rr} -3L_1: & -6x + 9y = 24 \\ 2L_2: & 6x + 8y = 10 \\ \hline \text{Addition}: & 17y = 34 \end{array}$$

We now solve the new equation for y, obtaining $y = 2$. We substitute $y = 2$ into one of the original equations, say L_1, and solve for the other unknown x, obtaining

$$2x - 3(2) = -8 \qquad \text{or} \qquad 2x - 6 = 8 \qquad \text{or} \qquad 2x = -2 \qquad \text{or} \qquad x = -1$$

Thus, $x = -1$, $y = 2$, or the pair $u = (-1, 2)$ is the unique solution of the system. The unique solution is expected, because $2/3 \neq -3/4$. [Geometrically, the lines corresponding to the equations intersect at the point $(-1, 2)$.]

EXAMPLE 3.6 (Nonunique Cases)

(a) Solve the system

$$\begin{aligned} L_1: &\quad x - 3y = 4 \\ L_2: &\quad -2x + 6y = 5 \end{aligned}$$

We eliminated x from the equations by multiplying L_1 by 2 and adding it to L_2—that is, by forming the new equation $L = 2L_1 + L_2$. This yields the degenerate equation

$$0x + 0y = 13$$

which has a nonzero constant $b = 13$. Thus, this equation and the system have no solution. This is expected, because $1/(-2) = -3/6 \neq 4/5$. (Geometrically, the lines corresponding to the equations are parallel.)

(b) Solve the system

$$\begin{aligned} L_1: &\quad x - 3y = 4 \\ L_2: &\quad -2x + 6y = -8 \end{aligned}$$

We eliminated x from the equations by multiplying L_1 by 2 and adding it to L_2—that is, by forming the new equation $L = 2L_1 + L_2$. This yields the degenerate equation

$$0x + 0y = 0$$

where the constant term is also zero. Thus, the system has an infinite number of solutions, which correspond to the solutions of either equation. This is expected, because $1/(-2) = -3/6 = 4/(-8)$. (Geometrically, the lines corresponding to the equations coincide.)

To find the general solution, let $y = a$, and substitute into L_1 to obtain

$$x - 3a = 4 \qquad \text{or} \qquad x = 3a + 4$$

Thus, the general solution of the system is

$$x = 3a + 4, y = a \qquad \text{or} \qquad u = (3a + 4,\ a)$$

where a (called a *parameter*) is any scalar.

3.5 Systems in Triangular and Echelon Forms

The main method for solving systems of linear equations, Gaussian elimination, is treated in Section 3.6. Here we consider two simple types of systems of linear equations: systems in triangular form and the more general systems in echelon form.

Triangular Form

Consider the following system of linear equations, which is in *triangular form*:

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 - 2x_4 &= 9 \\ 5x_2 - x_3 + 3x_4 &= 1 \\ 7x_3 - x_4 &= 3 \\ 2x_4 &= 8 \end{aligned}$$

That is, the first unknown x_1 is the leading unknown in the first equation, the second unknown x_2 is the leading unknown in the second equation, and so on. Thus, in particular, the system is square and each leading unknown is *directly* to the right of the leading unknown in the preceding equation.

Such a triangular system always has a unique solution, which may be obtained by *back-substitution*. That is,

(1) First solve the last equation for the last unknown to get $x_4 = 4$.

(2) Then substitute this value $x_4 = 4$ in the next-to-last equation, and solve for the next-to-last unknown x_3 as follows:

$$7x_3 - 4 = 3 \quad \text{or} \quad 7x_3 = 7 \quad \text{or} \quad x_3 = 1$$

(3) Now substitute $x_3 = 1$ and $x_4 = 4$ in the second equation, and solve for the second unknown x_2 as follows:

$$5x_2 - 1 + 12 = 1 \quad \text{or} \quad 5x_2 + 11 = 1 \quad \text{or} \quad 5x_2 = -10 \quad \text{or} \quad x_2 = -2$$

(4) Finally, substitute $x_2 = -2$, $x_3 = 1$, $x_4 = 4$ in the first equation, and solve for the first unknown x_1 as follows:

$$2x_1 + 6 + 5 - 8 = 9 \quad \text{or} \quad 2x_1 + 3 = 9 \quad \text{or} \quad 2x_1 = 6 \quad \text{or} \quad x_1 = 3$$

Thus, $x_1 = 3$, $x_2 = -2$, $x_3 = 1$, $x_4 = 4$, or, equivalently, the vector $u = (3, -2, 1, 4)$ is the unique solution of the system.

Remark: There is an alternative form for back-substitution (which will be used when solving a system using the matrix format). Namely, after first finding the value of the last unknown, we substitute this value for the last unknown in all the preceding equations before solving for the next-to-last unknown. This yields a triangular system with one less equation and one less unknown. For example, in the above triangular system, we substitute $x_4 = 4$ in all the preceding equations to obtain the triangular system

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 &= 17 \\ 5x_2 - x_3 &= -1 \\ 7x_3 &= 7 \end{aligned}$$

We then repeat the process using the new last equation. And so on.

Echelon Form, Pivot and Free Variables

The following system of linear equations is said to be in *echelon form*:

$$\begin{aligned} 2x_1 + 6x_2 - x_3 + 4x_4 - 2x_5 &= 15 \\ x_3 + 2x_4 + 2x_5 &= 5 \\ 3x_4 - 9x_5 &= 6 \end{aligned}$$

That is, no equation is degenerate and the leading unknown in each equation other than the first is to the right of the leading unknown in the preceding equation. The leading unknowns in the system, x_1, x_3, x_4, are called *pivot* variables, and the other unknowns, x_2 and x_5, are called *free* variables.

Generally speaking, an *echelon system* or a *system in echelon form* has the following form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + \cdots + a_{1n}x_n &= b_1 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{rj_r}x_{j_r} + \cdots + a_{rn}x_n &= b_r \end{aligned} \tag{3.5}$$

where $1 < j_2 < \cdots < j_r$ and a_{11}, $a_{2j_2}, \ldots, a_{rj_r}$ are not zero. The *pivot* variables are x_1, $x_{j_2}, \ldots, x_{j_r}$. Note that $r \leq n$.

The solution set of any echelon system is described in the following theorem (proved in Problem 3.10).

THEOREM 3.6: Consider a system of linear equations in echelon form, say with r equations in n unknowns. There are two cases:

(i) $r = n$. That is, there are as many equations as unknowns (triangular form). Then the system has a unique solution.

(ii) $r < n$. That is, there are more unknowns than equations. Then we can arbitrarily assign values to the $n - r$ free variables and solve uniquely for the r pivot variables, obtaining a solution of the system.

Suppose an echelon system contains more unknowns than equations. Assuming the field K is infinite, the system has an infinite number of solutions, because each of the $n - r$ free variables may be assigned any scalar.

The general solution of a system with free variables may be described in either of two equivalent ways, which we illustrate using the above echelon system where there are $r = 3$ equations and $n = 5$ unknowns. One description is called the ''Parametric Form'' of the solution, and the other description is called the ''Free-Variable Form.''

Parametric Form

Assign arbitrary values, called *parameters*, to the free variables x_2 and x_5, say $x_2 = a$ and $x_5 = b$, and then use back-substitution to obtain values for the pivot variables x_1, x_3, x_5 in terms of the parameters a and b. Specifically,

(1) Substitute $x_5 = b$ in the last equation, and solve for x_4:

$$3x_4 - 9b = 6 \qquad \text{or} \qquad 3x_4 = 6 + 9b \qquad \text{or} \qquad x_4 = 2 + 3b$$

(2) Substitute $x_4 = 2 + 3b$ and $x_5 = b$ into the second equation, and solve for x_3:

$$x_3 + 2(2 + 3b) + 2b = 5 \qquad \text{or} \qquad x_3 + 4 + 8b = 5 \qquad \text{or} \qquad x_3 = 1 - 8b$$

(3) Substitute $x_2 = a$, $x_3 = 1 - 8b$, $x_4 = 2 + 3b$, $x_5 = b$ into the first equation, and solve for x_1:

$$2x_1 + 6a - (1 - 8b) + 4(2 + 3b) - 2b = 15 \qquad \text{or} \qquad x_1 = 4 - 3a - 9b$$

Accordingly, the general solution in *parametric form* is

$$x_1 = 4 - 3a - 9b, \qquad x_2 = a, \qquad x_3 = 1 - 8b, \qquad x_4 = 2 + 3b, \qquad x_5 = b$$

or, equivalently, $v = (4 - 3a - 9b,\ a,\ 1 - 8b,\ 2 + 3b,\ b)$ where a and b are arbitrary numbers.

Free-Variable Form

Use back-substitution to solve for the pivot variables x_1, x_3, x_4 directly in terms of the free variables x_2 and x_5. That is, the last equation gives $x_4 = 2 + 3x_5$. Substitution in the second equation yields $x_3 = 1 - 8x_5$, and then substitution in the first equation yields $x_1 = 4 - 3x_2 - 9x_5$. Accordingly,

$$x_1 = 4 - 3x_2 - 9x_5, \quad x_2 = \text{free variable}, \quad x_3 = 1 - 8x_5, \quad x_4 = 2 + 3x_5, \quad x_5 = \text{free variable}$$

or, equivalently,

$$v = (4 - 3x_2 - 9x_5,\ x_2,\ 1 - 8x_5,\ 2 + 3x_5,\ x_5)$$

is the *free-variable form* for the general solution of the system.

We emphasize that there is no difference between the above two forms of the general solution, and the use of one or the other to represent the general solution is simply a matter of taste.

Remark: A particular solution of the above system can be found by assigning any values to the free variables and then solving for the pivot variables by back-substitution. For example, setting $x_2 = 1$ and $x_5 = 1$, we obtain

$$x_4 = 2 + 3 = 5, \qquad x_3 = 1 - 8 = -7, \qquad x_1 = 4 - 3 - 9 = -8$$

Thus, $u = (-8, 1, 7, 5, 1)$ is the particular solution corresponding to $x_2 = 1$ and $x_5 = 1$.

3.6 Gaussian Elimination

The main method for solving the general system (3.2) of linear equations is called *Gaussian elimination*. It essentially consists of two parts:

Part A. (Forward Elimination) Step-by-step reduction of the system yielding either a degenerate equation with no solution (which indicates the system has no solution) or an equivalent simpler system in triangular or echelon form.

Part B. (Backward Elimination) Step-by-step back-substitution to find the solution of the simpler system.

Part B has already been investigated in Section 3.4. Accordingly, we need only give the algorithm for Part A, which is as follows.

ALGORITHM 3.2 for (Part A): Input: The $m \times n$ system (3.2) of linear equations.

ELIMINATION STEP: Find the first unknown in the system with a nonzero coefficient (which now must be x_1).

(a) Arrange so that $a_{11} \neq 0$. That is, if necessary, interchange equations so that the first unknown x_1 appears with a nonzero coefficient in the first equation.

(b) Use a_{11} *as a pivot* to eliminate x_1 from all equations except the first equation. That is, for $i > 1$:

(1) Set $m = -a_{i1}/a_{11}$; (2) Replace L_i by $mL_1 + L_i$

The system now has the following form:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
a_{2j_2}x_{j_2} + \cdots + a_{2n}x_n &= b_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
a_{mj_2}x_{j_2} + \cdots + a_{mn}x_n &= b_n
\end{aligned}$$

where x_1 does not appear in any equation except the first, $a_{11} \neq 0$, and x_{j_2} denotes the first unknown with a nonzero coefficient in any equation other than the first.

(c) Examine each new equation L.

(1) If L has the form $0x_1 + 0x_2 + \cdots + 0x_n = b$ with $b \neq 0$, then

STOP

The system is *inconsistent* and has no solution.

(2) If L has the form $0x_1 + 0x_2 + \cdots + 0x_n = 0$ or if L is a multiple of another equation, then delete L from the system.

RECURSION STEP: Repeat the Elimination Step with each new ''smaller'' subsystem formed by all the equations excluding the first equation.

OUTPUT: Finally, the system is reduced to triangular or echelon form, or a degenerate equation with no solution is obtained indicating an inconsistent system.

The next remarks refer to the Elimination Step in Algorithm 3.2.

(1) The following number m in (b) is called the *multiplier*:

$$m = -\frac{a_{i1}}{a_{11}} = -\frac{\text{coefficient to be deleted}}{\text{pivot}}$$

(2) One could alternatively apply the following operation in (b):

Replace L_i by $-a_{i1}L_1 + a_{11}L_i$

This would avoid fractions if all the scalars were originally integers.

Gaussian Elimination Example

Here we illustrate in detail Gaussian elimination using the following system of linear equations:

$$\begin{aligned} L_1&: & x - 3y - 2z &= 6 \\ L_2&: & 2x - 4y - 3z &= 8 \\ L_3&: & -3x + 6y + 8z &= -5 \end{aligned}$$

Part A. We use the coefficient 1 of x in the first equation L_1 as the pivot in order to eliminate x from the second equation L_2 and from the third equation L_3. This is accomplished as follows:

(1) Multiply L_1 by the multiplier $m = -2$ and add it to L_2; that is, "Replace L_2 by $-2L_1 + L_2$."

(2) Multiply L_1 by the multiplier $m = 3$ and add it to L_3; that is, "Replace L_3 by $3L_1 + L_3$."

These steps yield

$$\begin{array}{rr} (-2)L_1: & -2x + 6y + 4z = -12 \\ L_2: & 2x - 4y - 3z = 8 \\ \hline \text{New } L_2: & 2y + z = -4 \end{array} \qquad \begin{array}{rr} 3L_1: & 3x - 9y - 6z = 18 \\ L_3: & -3x + 6y + 8z = -5 \\ \hline \text{New } L_3: & -3y + 2z = 13 \end{array}$$

Thus, the original system is replaced by the following system:

$$\begin{aligned} L_1&: & x - 3y - 2z &= 6 \\ L_2&: & 2y + z &= -4 \\ L_3&: & -3y + 2z &= 13 \end{aligned}$$

(Note that the equations L_2 and L_3 form a subsystem with one less equation and one less unknown than the original system.)

Next we use the coefficient 2 of y in the (new) second equation L_2 as the pivot in order to eliminate y from the (new) third equation L_3. This is accomplished as follows:

(3) Multiply L_2 by the multiplier $m = \frac{3}{2}$ and add it to L_3; that is, "Replace L_3 by $\frac{3}{2}L_2 + L_3$." (Alternately, "Replace L_3 by $3L_2 + 2L_3$," which will avoid fractions.)

This step yields

$$\begin{array}{rr} \frac{3}{2}L_2: & 3y + \frac{3}{2}z = -6 \\ L_3: & -3y + 2z = 13 \\ \hline \text{New } L_3: & \frac{7}{2}z = 7 \end{array} \qquad \text{or} \qquad \begin{array}{rr} 3L_2: & 6y + 3z = -12 \\ 2L_3: & -6y + 4z = 26 \\ \hline \text{New } L_3: & 7z = 14 \end{array}$$

Thus, our system is replaced by the following system:

$$\begin{aligned} L_1&: & x - 3y - 2z &= 6 \\ L_2&: & 2y + z &= -4 \\ L_3&: & 7z &= 14 \quad (\text{or } \tfrac{7}{2}z = 7) \end{aligned}$$

The system is now in triangular form, so Part A is completed.

Part B. The values for the unknowns are obtained in reverse order, z, y, x, by back-substitution. Specifically,

(1) Solve for z in L_3 to get $z = 2$.

(2) Substitute $z = 2$ in L_2, and solve for y to get $y = -3$.

(3) Substitute $y = -3$ and $z = 2$ in L_1, and solve for x to get $x = 1$.

Thus, the solution of the triangular system and hence the original system is as follows:

$$x = 1, \quad y = -3, \quad z = 2 \qquad \text{or, equivalently,} \qquad u = (1, -3, 2).$$

Condensed Format

The Gaussian elimination algorithm involves rewriting systems of linear equations. Sometimes we can avoid excessive recopying of some of the equations by adopting a ‘‘condensed format.’’ This format for the solution of the above system follows:

Number	Equation	Operation
(1)	$x - 3y - 2z = 6$	
(2)	$2x - 4y - 3z = 8$	
(3)	$-3x + 6y + 8z = -5$	
(2′)	$2y + z = -4$	Replace L_2 by $-2L_1 + L_2$
(3′)	$-3y + 2z = 13$	Replace L_3 by $3L_1 + L_3$
(3″)	$7z = 14$	Replace L_3 by $3L_2 + 2L_3$

That is, first we write down the number of each of the original equations. As we apply the Gaussian elimination algorithm to the system, we only write down the new equations, and we label each new equation using the same number as the original corresponding equation, but with an added prime. (After each new equation, we will indicate, for instructional purposes, the elementary operation that yielded the new equation.)

The system in triangular form consists of equations (1), (2′), and (3″), the numbers with the largest number of primes. Applying back-substitution to these equations again yields $x = 1$, $y = -3$, $z = 2$.

Remark: If two equations need to be interchanged, say to obtain a nonzero coefficient as a pivot, then this is easily accomplished in the format by simply renumbering the two equations rather than changing their positions.

EXAMPLE 3.7 Solve the following system:

$$\begin{aligned} x + 2y - 3z &= 1 \\ 2x + 5y - 8z &= 4 \\ 3x + 8y - 13z &= 7 \end{aligned}$$

We solve the system by Gaussian elimination.

Part A. (Forward Elimination) We use the coefficient 1 of x in the first equation L_1 as the pivot in order to eliminate x from the second equation L_2 and from the third equation L_3. This is accomplished as follows:

(1) Multiply L_1 by the multiplier $m = -2$ and add it to L_2; that is, ‘‘Replace L_2 by $-2L_1 + L_2$.’’

(2) Multiply L_1 by the multiplier $m = -3$ and add it to L_3; that is, ‘‘Replace L_3 by $-3L_1 + L_3$.’’

The two steps yield

$$\begin{aligned} x + 2y - 3z &= 1 \\ y - 2z &= 2 \\ 2y - 4z &= 4 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + 2y - 3z &= 1 \\ y - 2z &= 2 \end{aligned}$$

(The third equation is deleted, because it is a multiple of the second equation.) The system is now in echelon form with free variable z.

Part B. (Backward Elimination) To obtain the general solution, let the free variable $z = a$, and solve for x and y by back-substitution. Substitute $z = a$ in the second equation to obtain $y = 2 + 2a$. Then substitute $z = a$ and $y = 2 + 2a$ into the first equation to obtain

$$x + 2(2 + 2a) - 3a = 1 \qquad \text{or} \qquad x + 4 + 4a - 3a = 1 \qquad \text{or} \qquad x = -3 - a$$

Thus, the following is the general solution where a is a parameter:

$$x = -3 - a, \quad y = 2 + 2a, \quad z = a \qquad \text{or} \qquad u = (-3 - a,\ 2 + 2a,\ a)$$

EXAMPLE 3.8 Solve the following system:

$$\begin{aligned} x_1 + 3x_2 - \ 2x_3 + \ 5x_4 &= 4 \\ 2x_1 + 8x_2 - \ \ x_3 + \ 9x_4 &= 9 \\ 3x_1 + 5x_2 - 12x_3 + 17x_4 &= 7 \end{aligned}$$

We use Gaussian elimination.

Part A. (Forward Elimination) We use the coefficient 1 of x_1 in the first equation L_1 as the pivot in order to eliminate x_1 from the second equation L_2 and from the third equation L_3. This is accomplished by the following operations:

(1) ''Replace L_2 by $-2L_1 + L_2$'' and (2) ''Replace L_3 by $-3L_1 + L_3$''

These yield:

$$\begin{aligned} x_1 + 3x_2 - 2x_3 + 5x_4 &= \ \ 4 \\ 2x_2 + 3x_3 - \ x_4 &= \ \ 1 \\ -4x_2 - 6x_3 + 2x_4 &= -5 \end{aligned}$$

We now use the coefficient 2 of x_2 in the second equation L_2 as the pivot and the multiplier $m = 2$ in order to eliminate x_2 from the third equation L_3. This is accomplished by the operation ''Replace L_3 by $2L_2 + L_3$,'' which then yields the degenerate equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -3$$

This equation and, hence, the original system have no solution:

DO NOT CONTINUE

Remark 1: As in the above examples, Part A of Gaussian elimination tells us whether or not the system has a solution—that is, whether or not the system is consistent. Accordingly, Part B need never be applied when a system has no solution.

Remark 2: If a system of linear equations has more than four unknowns and four equations, then it may be more convenient to use the matrix format for solving the system. This matrix format is discussed later.

3.7 Echelon Matrices, Row Canonical Form, Row Equivalence

One way to solve a system of linear equations is by working with its augmented matrix M rather than the system itself. This section introduces the necessary matrix concepts for such a discussion. These concepts, such as echelon matrices and elementary row operations, are also of independent interest.

Echelon Matrices

A matrix A is called an *echelon matrix*, or is said to be in *echelon form*, if the following two conditions hold (where a *leading nonzero element* of a row of A is the first nonzero element in the row):

(1) All zero rows, if any, are at the bottom of the matrix.

(2) Each leading nonzero entry in a row is to the right of the leading nonzero entry in the preceding row.

That is, $A = [a_{ij}]$ is an echelon matrix if there exist nonzero entries

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}, \qquad \text{where} \quad j_1 < j_2 < \cdots < j_r$$

with the property that

$$a_{ij} = 0 \qquad \text{for} \quad \begin{cases} \text{(i) } i \le r, \quad j < j_i \\ \text{(ii) } i > r \end{cases}$$

The entries $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$, which are the leading nonzero elements in their respective rows, are called the *pivots* of the echelon matrix.

EXAMPLE 3.9 The following is an echelon matrix whose pivots have been circled:

$$A = \begin{bmatrix} 0 & \textcircled{2} & 3 & 4 & 5 & 9 & 0 & 7 \\ 0 & 0 & 0 & \textcircled{3} & 4 & 1 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{5} & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & \textcircled{8} & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Observe that the pivots are in columns C_2, C_4, C_6, C_7, and each is to the right of the one above. Using the above notation, the pivots are

$$a_{1j_1} = 2, \qquad a_{2j_2} = 3, \qquad a_{3j_3} = 5, \qquad a_{4j_4} = 8$$

where $j_1 = 2, \; j_2 = 4, \; j_3 = 6, \; j_4 = 7$. Here $r = 4$.

Row Canonical Form

A matrix A is said to be in *row canonical form* (or *row-reduced echelon form*) if it is an echelon matrix— that is, if it satisfies the above properties (1) and (2), and if it satisfies the following additional two properties:

(3) Each pivot (leading nonzero entry) is equal to 1.

(4) Each pivot is the only nonzero entry in its column.

The major difference between an echelon matrix and a matrix in row canonical form is that in an echelon matrix there must be zeros below the pivots [Properties (1) and (2)], but in a matrix in row canonical form, each pivot must also equal 1 [Property (3)] and there must also be zeros above the pivots [Property (4)].

The zero matrix 0 of any size and the identity matrix I of any size are important special examples of matrices in row canonical form.

EXAMPLE 3.10

The following are echelon matrices whose pivots have been circled:

$$\begin{bmatrix} \textcircled{2} & 3 & 2 & 0 & 4 & 5 & -6 \\ 0 & 0 & \textcircled{0} & \textcircled{1} & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{6} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} \textcircled{1} & 2 & 3 \\ 0 & 0 & \textcircled{1} \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & \textcircled{1} & 3 & 0 & 0 & 4 \\ 0 & 0 & 0 & \textcircled{1} & 0 & -3 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 2 \end{bmatrix}$$

The third matrix is also an example of a matrix in row canonical form. The second matrix is not in row canonical form, because it does not satisfy property (4); that is, there is a nonzero entry above the second pivot in the third column. The first matrix is not in row canonical form, because it satisfies neither property (3) nor property (4); that is, some pivots are not equal to 1 and there are nonzero entries above the pivots.

Elementary Row Operations

Suppose A is a matrix with rows $R_1, R_2, \ldots, R_m$. The following operations on A are called *elementary row operations*.

[E$_1$] (Row Interchange): Interchange rows R_i and R_j. This may be written as

$$\text{"Interchange } R_i \text{ and } R_j\text{"} \qquad \text{or} \qquad \text{"}R_i \longleftrightarrow R_j\text{"}$$

[E$_2$] (Row Scaling): Replace row R_i by a nonzero multiple kR_i of itself. This may be written as

$$\text{"Replace } R_i \text{ by } kR_i \ (k \neq 0)\text{"} \qquad \text{or} \qquad \text{"}kR_i \to R_i\text{"}$$

[E$_3$] (Row Addition): Replace row R_j by the sum of a multiple kR_i of a row R_i and itself. This may be written as

$$\text{"Replace } R_j \text{ by } kR_i + R_j\text{"} \qquad \text{or} \qquad \text{"}kR_i + R_j \to R_j\text{"}$$

The arrow $\to$ in E$_2$ and E$_3$ may be read as "replaces."

Sometimes (say to avoid fractions when all the given scalars are integers) we may apply [E$_2$] and [E$_3$] in one step; that is, we may apply the following operation:

[E] Replace R_j by the sum of a multiple kR_i of a row R_i and a nonzero multiple $k'R_j$ of itself. This may be written as

$$\text{"Replace } R_j \text{ by } kR_i + k'R_j \ (k' \neq 0)\text{"} \qquad \text{or} \qquad \text{"}kR_i + k'R_j \to R_j\text{"}$$

We emphasize that in operations [E$_3$] and [E] only row R_j is changed.

Row Equivalence, Rank of a Matrix

A matrix A is said to be *row equivalent* to a matrix B, written

$A \sim B$

if B can be obtained from A by a sequence of elementary row operations. In the case that B is also an echelon matrix, B is called an *echelon form* of A.

The following are two basic results on row equivalence.

THEOREM 3.7: Suppose $A = [a_{ij}]$ and $B = [b_{ij}]$ are row equivalent echelon matrices with respective pivot entries

$$a_{1j_1}, a_{2j_2}, \ldots a_{rj_r} \qquad \text{and} \qquad b_{1k_1}, b_{2k_2}, \ldots b_{sk_s}$$

Then A and B have the same number of nonzero rows—that is, $r = s$—and the pivot entries are in the same positions—that is, $j_1 = k_1, \quad j_2 = k_2, \quad \ldots, \quad j_r = k_r$.

THEOREM 3.8: Every matrix A is row equivalent to a unique matrix in row canonical form.

The proofs of the above theorems will be postponed to Chapter 4. The unique matrix in Theorem 3.8 is called the *row canonical form* of A.

Using the above theorems, we can now give our first definition of the rank of a matrix.

DEFINITION: The *rank* of a matrix A, written rank(A), is equal to the number of pivots in an echelon form of A.

> The rank is a very important property of a matrix and, depending on the context in which the matrix is used, it will be defined in many different ways. Of course, all the definitions lead to the same number.

The next section gives the matrix format of Gaussian elimination, which finds an echelon form of any matrix A (and hence the rank of A), and also finds the row canonical form of A.