One can show that row equivalence is an *equivalence relation*. That is,

(1) $A \sim A$ for any matrix A.
(2) If $A \sim B$, then $B \sim A$.
(3) If $A \sim B$ and $B \sim C$, then $A \sim C$.

Property (2) comes from the fact that each elementary row operation has an inverse operation of the same type. Namely,

(i) ‘‘Interchange R_i and R_j’’ is its own inverse.
(ii) ‘‘Replace R_i by kR_i’’ and ‘‘Replace R_i by $(1/k)R_i$’’ are inverses.
(iii) ‘‘Replace R_j by $kR_i + R_j$’’ and ‘‘Replace R_j by $-kR_i + R_j$’’ are inverses.

There is a similar result for operation [E] (Problem 3.73).

3.8 Gaussian Elimination, Matrix Formulation

This section gives two matrix algorithms that accomplish the following:

(1) Algorithm 3.3 transforms any matrix A into an echelon form.
(2) Algorithm 3.4 transforms the echelon matrix into its row canonical form.

These algorithms, which use the elementary row operations, are simply restatements of *Gaussian elimination* as applied to matrices rather than to linear equations. (The term ‘‘row reduce’’ or simply ‘‘reduce’’ will mean to transform a matrix by the elementary row operations.)

ALGORITHM 3.3 (Forward Elimination): The input is any matrix A. (The algorithm puts 0’s below each pivot, working from the ‘‘top-down.’’) The output is an echelon form of A.

Step 1. Find the first column with a nonzero entry. Let j_1 denote this column.

(a) Arrange so that $a_{1j_1} \neq 0$. That is, if necessary, interchange rows so that a nonzero entry appears in the first row in column j_1.

(b) Use a_{1j_1} as a pivot to obtain 0’s below a_{1j_1}.

Specifically, for $i > 1$:

$$(1)\ \text{Set } m = -a_{ij_1}/a_{1j_1}; \qquad (2)\ \text{Replace } R_i \text{ by } mR_1 + R_i$$

[That is, apply the operation $-(a_{ij_1}/a_{1j_1})R_1 + R_i \to R_i$.]

Step 2. Repeat Step 1 with the submatrix formed by all the rows excluding the first row. Here we let j_2 denote the first column in the subsystem with a nonzero entry. Hence, at the end of Step 2, we have $a_{2j_2} \neq 0$.

Steps 3 to *r*. Continue the above process until a submatrix has only zero rows.

We emphasize that at the end of the algorithm, the pivots will be

$$a_{1j_1},\ a_{2j_2},\ \ldots,\ a_{rj_r}$$

where r denotes the number of nonzero rows in the final echelon matrix.

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

$$m = -\frac{a_{ij_1}}{a_{1j_1}} = -\frac{\text{entry to be deleted}}{\text{pivot}}$$

Remark 2: One could replace the operation in Step 1(b) by the following which would avoid fractions if all the scalars were originally integers.

Replace R_i by $-a_{ij_1}R_1 + a_{1j_1}R_i$.

ALGORITHM 3.4 (Backward Elimination): The input is a matrix $A = [a_{ij}]$ in echelon form with pivot entries

$$a_{1j_1}, \quad a_{2j_2}, \quad \ldots, \quad a_{rj_r}$$

The output is the row canonical form of A.

Step 1. (a) (Use row scaling so the last pivot equals 1.) Multiply the last nonzero row R_r by $1/a_{rj_r}$.

(b) (Use $a_{rj_r} = 1$ to obtain 0's above the pivot.) For $i = r-1, \quad r-2, \quad \ldots, \quad 2, \quad 1$:

$$(1)\ \text{Set } m = -a_{ij_r}; \qquad (2)\ \text{Replace } R_i \text{ by } mR_r + R_i$$

(That is, apply the operations $-a_{ij_r}R_r + R_i \to R_i$.)

Steps 2 to *r*−1. Repeat Step 1 for rows $R_{r-1}, R_{r-2}, \ldots, R_2$.

Step *r*. (Use row scaling so the first pivot equals 1.) Multiply R_1 by $1/a_{1j_1}$.

There is an alternative form of Algorithm 3.4, which we describe here in words. The formal description of this algorithm is left to the reader as a supplementary problem.

ALTERNATIVE ALGORITHM 3.4 Puts 0's above the pivots row by row from the bottom up (rather than column by column from right to left).

The alternative algorithm, when applied to an augmented matrix M of a system of linear equations, is essentially the same as solving for the pivot unknowns one after the other from the bottom up.

Remark: We emphasize that Gaussian elimination is a two-stage process. Specifically,

Stage A (Algorithm 3.3). Puts 0's below each pivot, working from the top row R_1 down.

Stage B (Algorithm 3.4). Puts 0's above each pivot, working from the bottom row R_r up.

There is another algorithm, called *Gauss–Jordan*, that also row reduces a matrix to its row canonical form. The difference is that Gauss–Jordan puts 0's both below and above each pivot as it works its way from the top row R_1 down. Although Gauss–Jordan may be easier to state and understand, it is much less efficient than the two-stage Gaussian elimination algorithm.

EXAMPLE 3.11 Consider the matrix $A = \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 2 & 4 & -4 & 6 & 10 \\ 3 & 6 & -6 & 9 & 13 \end{bmatrix}$.

(a) Use Algorithm 3.3 to reduce A to an echelon form.

(b) Use Algorithm 3.4 to further reduce A to its row canonical form.

(a) First use $a_{11} = 1$ as a pivot to obtain 0's below a_{11}; that is, apply the operations ''Replace R_2 by $-2R_1 + R_2$'' and ''Replace R_3 by $-3R_1 + R_3$.'' Then use $a_{23} = 2$ as a pivot to obtain 0 below a_{23}; that is, apply the operation ''Replace R_3 by $-\frac{3}{2}R_2 + R_3$.'' This yields

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

The matrix is now in echelon form.

(b) Multiply R_3 by $-\frac{1}{2}$ so the pivot entry $a_{35} = 1$, and then use $a_{35} = 1$ as a pivot to obtain 0's above it by the operations "Replace R_2 by $-6R_3 + R_2$" and then "Replace R_1 by $-2R_3 + R_1$." This yields

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

Multiply R_2 by $\frac{1}{2}$ so the pivot entry $a_{23} = 1$, and then use $a_{23} = 1$ as a pivot to obtain 0's above it by the operation "Replace R_1 by $3R_2 + R_1$." This yields

$$A \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 7 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The last matrix is the row canonical form of A.

Application to Systems of Linear Equations

One way to solve a system of linear equations is by working with its augmented matrix M rather than the equations themselves. Specifically, we reduce M to echelon form (which tells us whether the system has a solution), and then further reduce M to its row canonical form (which essentially gives the solution of the original system of linear equations). The justification for this process comes from the following facts:

(1) Any elementary row operation on the augmented matrix M of the system is equivalent to applying the corresponding operation on the system itself.

(2) The system has a solution if and only if the echelon form of the augmented matrix M does not have a row of the form $(0, 0, \ldots, 0, b)$ with $b \neq 0$.

(3) In the row canonical form of the augmented matrix M (excluding zero rows), the coefficient of each basic variable is a pivot entry equal to 1, and it is the only nonzero entry in its respective column; hence, the free-variable form of the solution of the system of linear equations is obtained by simply transferring the free variables to the other side.

This process is illustrated below.

EXAMPLE 3.12 Solve each of the following systems:

$$\begin{aligned} x_1 + x_2 - 2x_3 + 4x_4 &= 5 \\ 2x_1 + 2x_2 - 3x_3 + x_4 &= 3 \\ 3x_1 + 3x_2 - 4x_3 - 2x_4 &= 1 \end{aligned} \qquad \begin{aligned} x_1 + x_2 - 2x_3 + 3x_4 &= 4 \\ 2x_1 + 3x_2 + 3x_3 - x_4 &= 3 \\ 5x_1 + 7x_2 + 4x_3 + x_4 &= 5 \end{aligned} \qquad \begin{aligned} x + 2y + z &= 3 \\ 2x + 5y - z &= -4 \\ 3x - 2y - z &= 5 \end{aligned}$$

(a) (b) (c)

(a) Reduce its augmented matrix M to echelon form and then to row canonical form as follows:

$$M = \begin{bmatrix} 1 & 1 & -2 & 4 & 5 \\ 2 & 2 & -3 & 1 & 3 \\ 3 & 3 & -4 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 4 & 5 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 2 & -14 & -14 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & -10 & -9 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rewrite the row canonical form in terms of a system of linear equations to obtain the free variable form of the solution. That is,

$$\begin{aligned} x_1 + x_2 \quad - 10x_4 &= -9 \\ x_3 - 7x_4 &= -7 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x_1 &= -9 - x_2 + 10x_4 \\ x_3 &= -7 + 7x_4 \end{aligned}$$

(The zero row is omitted in the solution.) Observe that x_1 and x_3 are the pivot variables, and x_2 and x_4 are the free variables.

(b) First reduce its augmented matrix M to echelon form as follows:

$$M = \begin{bmatrix} 1 & 1 & -2 & 3 & 4 \\ 2 & 3 & 3 & -1 & 3 \\ 5 & 7 & 4 & 1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 2 & 14 & -14 & -15 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 0 & 0 & 0 & -5 \end{bmatrix}$$

There is no need to continue to find the row canonical form of M, because the echelon form already tells us that the system has no solution. Specifically, the third row of the echelon matrix corresponds to the degenerate equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -5$$

which has no solution. Thus, the system has no solution.

(c) Reduce its augmented matrix M to echelon form and then to row canonical form as follows:

$$M = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 5 & -1 & -4 \\ 3 & -2 & -1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & -8 & -4 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & 0 & -28 & -84 \end{bmatrix}$$

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

Thus, the system has the unique solution $x = 2$, $y = -1$, $z = 3$, or, equivalently, the vector $u = (2, -1, 3)$. We note that the echelon form of M already indicated that the solution was unique, because it corresponded to a triangular system.

Application to Existence and Uniqueness Theorems

This subsection gives theoretical conditions for the existence and uniqueness of a solution of a system of linear equations using the notion of the rank of a matrix.

THEOREM 3.9: Consider a system of linear equations in n unknowns with augmented matrix $M = [A, B]$. Then,

(a) The system has a solution if and only if $\text{rank}(A) = \text{rank}(M)$.

(b) The solution is unique if and only if $\text{rank}(A) = \text{rank}(M) = n$.

Proof of (a). The system has a solution if and only if an echelon form of $M = [A, B]$ does not have a row of the form

$$(0, 0, \ldots, 0, b), \qquad \text{with} \quad b \neq 0$$

If an echelon form of M does have such a row, then b is a pivot of M but not of A, and hence, $\text{rank}(M) > \text{rank}(A)$. Otherwise, the echelon forms of A and M have the same pivots, and hence, $\text{rank}(A) = \text{rank}(M)$. This proves (a).

Proof of (b). The system has a unique solution if and only if an echelon form has no free variable. This means there is a pivot for each unknown. Accordingly, $n = \text{rank}(A) = \text{rank}(M)$. This proves (b).

The above proof uses the fact (Problem 3.74) that an echelon form of the augmented matrix $M = [A, B]$ also automatically yields an echelon form of A.

3.9 Matrix Equation of a System of Linear Equations

The general system (3.2) of m linear equations in n unknowns is equivalent to the matrix equation

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix} \qquad \text{or} \qquad AX = B$$

where $A = [a_{ij}]$ is the coefficient matrix, $X = [x_j]$ is the column vector of unknowns, and $B = [b_i]$ is the column vector of constants. (Some texts write $Ax = b$ rather than $AX = B$, in order to emphasize that x and b are simply column vectors.)

The statement that the system of linear equations and the matrix equation are equivalent means that any vector solution of the system is a solution of the matrix equation, and vice versa.

EXAMPLE 3.13 The following system of linear equations and matrix equation are equivalent:

$$\begin{aligned} x_1 + 2x_2 - 4x_3 + 7x_4 &= 4 \\ 3x_1 - 5x_2 + 6x_3 - 8x_4 &= 8 \\ 4x_1 - 3x_2 - 2x_3 + 6x_4 &= 11 \end{aligned} \qquad \text{and} \qquad \begin{bmatrix} 1 & 2 & -4 & 7 \\ 3 & -5 & 6 & -8 \\ 4 & -3 & -2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 11 \end{bmatrix}$$

We note that $x_1 = 3,\ x_2 = 1,\ x_3 = 2,\ x_4 = 1$, or, in other words, the vector $u = [3, 1, 2, 1]$ is a solution of the system. Thus, the (column) vector u is also a solution of the matrix equation.

The matrix form $AX = B$ of a system of linear equations is notationally very convenient when discussing and proving properties of systems of linear equations. This is illustrated with our first theorem (described in Fig. 3-1), which we restate for easy reference.

THEOREM 3.1: Suppose the field K is infinite. Then the system $AX = B$ has: (a) a unique solution, (b) no solution, or (c) an infinite number of solutions.

Proof. It suffices to show that if $AX = B$ has more than one solution, then it has infinitely many. Suppose u and v are distinct solutions of $AX = B$; that is, $Au = B$ and $Av = B$. Then, for any $k \in K$,

$$A[u + k(u - v)] = Au + k(Au - Av) = B + k(B - B) = B$$

Thus, for each $k \in K$, the vector $u + k(u - v)$ is a solution of $AX = B$. Because all such solutions are distinct (Problem 3.47), $AX = B$ has an infinite number of solutions.

Observe that the above theorem is true when K is the real field $\mathbf{R}$ (or the complex field $\mathbf{C}$). Section 3.3 shows that the theorem has a geometrical description when the system consists of two equations in two unknowns, where each equation represents a line in $\mathbf{R}^2$. The theorem also has a geometrical description when the system consists of three nondegenerate equations in three unknowns, where the three equations correspond to planes H_1, H_2, H_3 in $\mathbf{R}^3$. That is,

(a) *Unique solution:* Here the three planes intersect in exactly one point.

(b) *No solution:* Here the planes may intersect pairwise but with no common point of intersection, or two of the planes may be parallel.

(c) *Infinite number of solutions:* Here the three planes may intersect in a line (one free variable), or they may coincide (two free variables).

These three cases are pictured in Fig. 3-3.

Matrix Equation of a Square System of Linear Equations

A system $AX = B$ of linear equations is square if and only if the matrix A of coefficients is square. In such a case, we have the following important result.

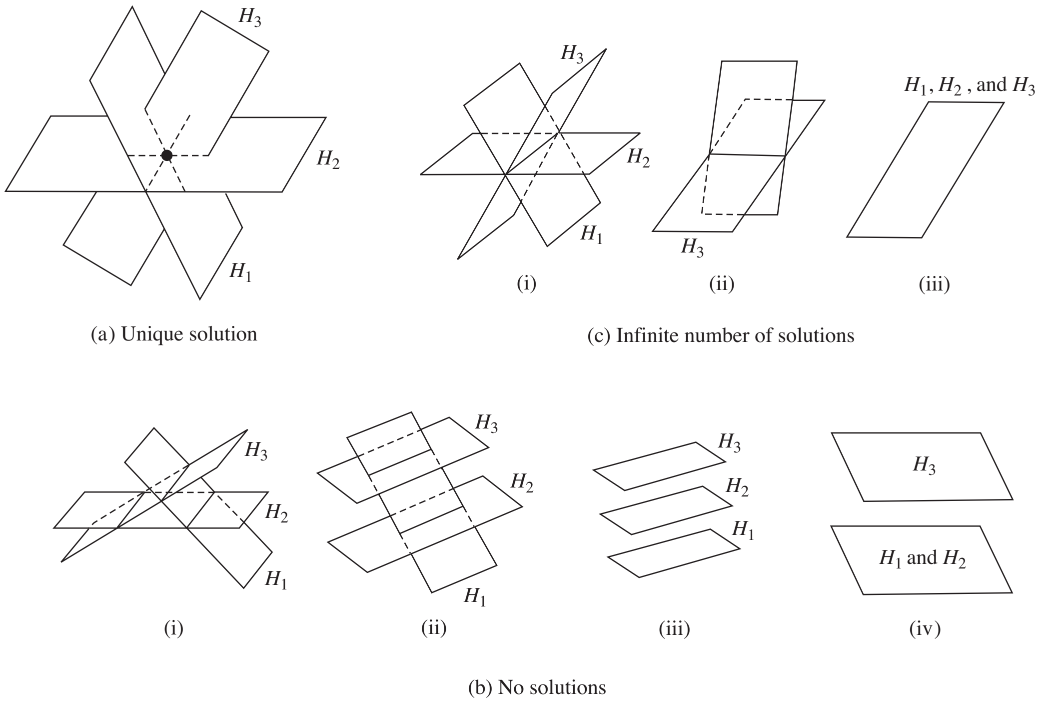

Figure 3-3

THEOREM 3.10: A square system $AX = B$ of linear equations has a unique solution if and only if the matrix A is invertible. In such a case, $A^{-1}B$ is the unique solution of the system.

We only prove here that if A is invertible, then $A^{-1}B$ is a unique solution. If A is invertible, then

$$A(A^{-1}B) = (AA^{-1})B = IB = B$$

and hence, $A^{-1}B$ is a solution. Now suppose v is any solution, so $Av = B$. Then

$$v = Iv = (A^{-1}A)v = A^{-1}(Av) = A^{-1}B$$

Thus, the solution $A^{-1}B$ is unique.

EXAMPLE 3.14 Consider the following system of linear equations, whose coefficient matrix A and inverse A^{-1} are also given:

$$\begin{aligned} x + 2y + \ \ 3z &= 1 \\ x + 3y + \ \ 6z &= 3 \\ 2x + 6y + 13z &= 5 \end{aligned}, \qquad A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 6 \\ 2 & 6 & 13 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} 3 & -8 & 3 \\ -1 & 7 & -3 \\ 0 & -2 & 1 \end{bmatrix}$$

By Theorem 3.10, the unique solution of the system is

$$A^{-1}B = \begin{bmatrix} 3 & -8 & 3 \\ -1 & 7 & -3 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} -6 \\ 5 \\ -1 \end{bmatrix}$$

That is, $x = -6$, $y = 5$, $z = -1$.

Remark: We emphasize that Theorem 3.10 does not usually help us to find the solution of a square system. That is, finding the inverse of a coefficient matrix A is not usually any easier than solving the system directly. Thus, unless we are given the inverse of a coefficient matrix A, as in Example 3.14, we usually solve a square system by Gaussian elimination (or some iterative method whose discussion lies beyond the scope of this text).

3.10 Systems of Linear Equations and Linear Combinations of Vectors

The general system (3.2) of linear equations may be rewritten as the following vector equation:

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

Recall that a vector v in K^n is said to be a *linear combination* of vectors $u_1, u_2, \dots, u_m$ in K^n if there exist scalars $a_1, a_2, \dots, a_m$ in K such that

$$v = a_1u_1 + a_2u_2 + \cdots + a_mu_m$$

Accordingly, the general system (3.2) of linear equations and the above equivalent vector equation have a solution if and only if the column vector of constants is a linear combination of the columns of the coefficient matrix. We state this observation formally.

THEOREM 3.11: A system $AX = B$ of linear equations has a solution if and only if B is a linear combination of the columns of the coefficient matrix A.

Thus, the answer to the problem of expressing a given vector v in K^n as a linear combination of vectors $u_1, u_2, \dots, u_m$ in K^n reduces to solving a system of linear equations.

Linear Combination Example

Suppose we want to write the vector $v = (1, -2, 5)$ as a linear combination of the vectors

$$u_1 = (1, 1, 1), \qquad u_2 = (1, 2, 3), \qquad u_3 = (2, -1, 1)$$

First we write $v = xu_1 + yu_2 + zu_3$ with unknowns x, y, z, and then we find the equivalent system of linear equations which we solve. Specifically, we first write

$$\begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + z \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \qquad (*)$$

Then

$$\begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} = \begin{bmatrix} x \\ x \\ x \end{bmatrix} + \begin{bmatrix} y \\ 2y \\ 3y \end{bmatrix} + \begin{bmatrix} 2z \\ -z \\ z \end{bmatrix} = \begin{bmatrix} x + y + 2z \\ x + 2y - z \\ x + 3y + z \end{bmatrix}$$

Setting corresponding entries equal to each other yields the following equivalent system:

$$\begin{aligned} x + \ \ y + 2z &= \ \ 1 \\ x + 2y - \ \ z &= -2 \\ x + 3y + \ \ z &= \ \ 5 \end{aligned} \qquad (**)$$

For notational convenience, we have written the vectors in $\mathbf{R}^n$ as columns, because it is then easier to find the equivalent system of linear equations. In fact, one can easily go from the vector equation (*) directly to the system (**).

Now we solve the equivalent system of linear equations by reducing the system to echelon form. This yields

$$\begin{aligned} x + \ \ y + 2z &= \ \ 1 \\ y - 3z &= -3 \\ 2y - \ \ z &= \ \ 4 \end{aligned} \qquad \text{and then} \qquad \begin{aligned} x + y + 2z &= \ \ 1 \\ y - 3z &= -3 \\ 5z &= 10 \end{aligned}$$

Back-substitution yields the solution $x = -6$, $y = 3$, $z = 2$. Thus, $v = -6u_1 + 3u_2 + 2u_3$.

EXAMPLE 3.15

(a) Write the vector $v = (4, 9, 19)$ as a linear combination of

$$u_1 = (1, -2, 3), \qquad u_2 = (3, -7, 10), \qquad u_3 = (2, 1, 9).$$

Find the equivalent system of linear equations by writing $v = xu_1 + yu_2 + zu_3$, and reduce the system to an echelon form. We have

$$\begin{array}{rl} x + 3y + 2z = & 4 \\ -2x - 7y + z = & 9 \\ 3x + 10y + 9z = & 19 \end{array} \quad \text{or} \quad \begin{array}{rl} x + 3y + 2z = & 4 \\ -y + 5z = & 17 \\ y + 3z = & 7 \end{array} \quad \text{or} \quad \begin{array}{rl} x + 3y + 2z = & 4 \\ -y + 5z = & 17 \\ 8z = & 24 \end{array}$$

Back-substitution yields the solution $x = 4$, $y = -2$, $z = 3$. Thus, v is a linear combination of u_1, u_2, u_3. Specifically, $v = 4u_1 - 2u_2 + 3u_3$.

(b) Write the vector $v = (2, 3, -5)$ as a linear combination of

$$u_1 = (1, 2, -3), \qquad u_2 = (2, 3, -4), \qquad u_3 = (1, 3, -5)$$

Find the equivalent system of linear equations by writing $v = xu_1 + yu_2 + zu_3$, and reduce the system to an echelon form. We have

$$\begin{array}{rl} x + 2y + z = & 2 \\ 2x + 3y + 3z = & 3 \\ -3x - 4y - 5z = & -5 \end{array} \quad \text{or} \quad \begin{array}{rl} x + 2y + z = & 2 \\ -y + z = & -1 \\ 2y - 2z = & 1 \end{array} \quad \text{or} \quad \begin{array}{rl} x + 2y + z = & 2 \\ -5y + 5z = & -1 \\ 0 = & 3 \end{array}$$

The system has no solution. Thus, it is impossible to write v as a linear combination of u_1, u_2, u_3.

Linear Combinations of Orthogonal Vectors, Fourier Coefficients

Recall first (Section 1.4) that the dot (inner) product $u \cdot v$ of vectors $u = (a_1, \ldots, a_n)$ and $v = (b_1, \ldots, b_n)$ in $\mathbf{R}^n$ is defined by

$$u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Furthermore, vectors u and v are said to be *orthogonal* if their dot product $u \cdot v = 0$.

Suppose that $u_1, u_2, \ldots, u_n$ in $\mathbf{R}^n$ are n nonzero pairwise orthogonal vectors. This means

$$\text{(i)} \quad u_i \cdot u_j = 0 \quad \text{for } i \neq j \qquad \text{and} \qquad \text{(ii)} \quad u_i \cdot u_i \neq 0 \quad \text{for each } i$$

Then, for any vector v in $\mathbf{R}^n$, there is an easy way to write v as a linear combination of $u_1, u_2, \ldots, u_n$, which is illustrated in the next example.

EXAMPLE 3.16 Consider the following three vectors in $\mathbf{R}^3$:

$$u_1 = (1, 1, 1), \qquad u_2 = (1, -3, 2), \qquad u_3 = (5, -1, -4)$$

These vectors are pairwise orthogonal; that is,

$$u_1 \cdot u_2 = 1 - 3 + 2 = 0, \qquad u_1 \cdot u_3 = 5 - 1 - 4 = 0, \qquad u_2 \cdot u_3 = 5 + 3 - 8 = 0$$

Suppose we want to write $v = (4, 14, -9)$ as a linear combination of u_1, u_2, u_3.

Method 1. Find the equivalent system of linear equations as in Example 3.14 and then solve, obtaining $v = 3u_1 - 4u_2 + u_3$.

Method 2. (This method uses the fact that the vectors u_1, u_2, u_3 are mutually orthogonal, and hence, the arithmetic is much simpler.) Set v as a linear combination of u_1, u_2, u_3 using unknown scalars x, y, z as follows:

$$(4, 14, -9) = x(1, 1, 1) + y(1, -3, 2) + z(5, -1, -4) \qquad (*)$$

Take the dot product of (*) with respect to u_1 to get

$$(4,14,-9)\cdot(1,1,1) = x(1,1,1)\cdot(1,1,1) \qquad \text{or} \qquad 9 = 3x \qquad \text{or} \qquad x = 3$$

(The last two terms drop out, because u_1 is orthogonal to u_2 and to u_3.) Next take the dot product of (*) with respect to u_2 to obtain

$$(4,14,-9)\cdot(1,-3,2) = y(1,-3,2)\cdot(1,-3,2) \qquad \text{or} \qquad -56 = 14y \qquad \text{or} \qquad y = -4$$

Finally, take the dot product of (*) with respect to u_3 to get

$$(4,14,-9)\cdot(5,-1,-4) = z(5,-1,-4)\cdot(5,-1,-4) \qquad \text{or} \qquad 42 = 42z \qquad \text{or} \qquad z = 1$$

Thus, $v = 3u_1 - 4u_2 + u_3$.

The procedure in Method 2 in Example 3.16 is valid in general. Namely,

THEOREM 3.12: Suppose $u_1, u_2, \ldots, u_n$ are nonzero mutually orthogonal vectors in $\mathbf{R}^n$. Then, for any vector v in $\mathbf{R}^n$,

$$v = \frac{v\cdot u_1}{u_1\cdot u_1}u_1 + \frac{v\cdot u_2}{u_2\cdot u_2}u_2 + \cdots + \frac{v\cdot u_n}{u_n\cdot u_n}u_n$$

We emphasize that there must be n such orthogonal vectors u_i in $\mathbf{R}^n$ for the formula to be used. Note also that each $u_i \cdot u_i \neq 0$, because each u_i is a nonzero vector.

Remark: The following scalar k_i (appearing in Theorem 3.12) is called the *Fourier coefficient* of v with respect to u_i:

$$k_i = \frac{v\cdot u_i}{u_i\cdot u_i} = \frac{v\cdot u_i}{\|u_i\|^2}$$

It is analogous to a coefficient in the celebrated Fourier series of a function.

3.11 Homogeneous Systems of Linear Equations

A system of linear equations is said to be *homogeneous* if all the constant terms are zero. Thus, a homogeneous system has the form $AX = 0$. Clearly, such a system always has the zero vector $0 = (0, 0, \ldots, 0)$ as a solution, called the *zero* or *trivial* solution. Accordingly, we are usually interested in whether or not the system has a nonzero solution.

Because a homogeneous system $AX = 0$ has at least the zero solution, it can always be put in an echelon form, say

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + \cdots + a_{1n}x_n &= 0 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{rj_r}x_{j_r} + \cdots + a_{rn}x_n &= 0 \end{aligned}$$

Here r denotes the number of equations in echelon form and n denotes the number of unknowns. Thus, the echelon system has $n - r$ free variables.

The question of nonzero solutions reduces to the following two cases:

(i) $r = n$. The system has only the zero solution.

(ii) $r < n$. The system has a nonzero solution.

Accordingly, if we begin with fewer equations than unknowns, then, in echelon form, $r < n$, and the system has a nonzero solution. This proves the following important result.

THEOREM 3.13: A homogeneous system $AX = 0$ with more unknowns than equations has a nonzero solution.

EXAMPLE 3.17 Determine whether or not each of the following homogeneous systems has a nonzero solution:

$$\begin{array}{ccc} \begin{aligned} x+\ y-\ z&=0\\ 2x-3y+\ z&=0\\ x-4y+2z&=0 \end{aligned} & \begin{aligned} x+\ y-\ z&=0\\ 2x+4y-\ z&=0\\ 3x+2y+2z&=0 \end{aligned} & \begin{aligned} x_1+2x_2-3x_3+4x_4&=0\\ 2x_1-3x_2+5x_3-7x_4&=0\\ 5x_1+6x_2-9x_3+8x_4&=0 \end{aligned} \\ \text{(a)} & \text{(b)} & \text{(c)} \end{array}$$

(a) Reduce the system to echelon form as follows:

$$\begin{aligned} x+\ y-\ z&=0\\ -5y+3z&=0\\ -5y+3z&=0 \end{aligned} \qquad \text{and then} \qquad \begin{aligned} x+\ y-\ z&=0\\ -5y+3z&=0 \end{aligned}$$

The system has a nonzero solution, because there are only two equations in the three unknowns in echelon form. Here z is a free variable. Let us, say, set $z=5$. Then, by back-substitution, $y=3$ and $x=2$. Thus, the vector $u=(2,3,5)$ is a particular nonzero solution.

(b) Reduce the system to echelon form as follows:

$$\begin{aligned} x+y-\ z&=0\\ 2y+\ z&=0\\ -y+5z&=0 \end{aligned} \qquad \text{and then} \qquad \begin{aligned} x+y-z&=0\\ 2y+z&=0\\ 11z&=0 \end{aligned}$$

In echelon form, there are three equations in three unknowns. Thus, the system has only the zero solution.

(c) The system must have a nonzero solution (Theorem 3.13), because there are four unknowns but only three equations. (Here we do not need to reduce the system to echelon form.)

Basis for the General Solution of a Homogeneous System

Let W denote the general solution of a homogeneous system $AX=0$. A list of nonzero solution vectors $u_1, u_2, \ldots, u_s$ of the system is said to be a *basis* for W if each solution vector $w \in W$ can be expressed uniquely as a linear combination of the vectors $u_1, u_2, \ldots, u_s$; that is, there exist unique scalars $a_1, a_2, \ldots, a_s$ such that

$$w = a_1u_1 + a_2u_2 + \cdots + a_su_s$$

The number s of such basis vectors is equal to the number of free variables. This number s is called the *dimension* of W, written as $\dim W = s$. When $W=\{0\}$—that is, the system has only the zero solution—we define $\dim W = 0$.

The following theorem, proved in Chapter 5, page 171, tells us how to find such a basis.

THEOREM 3.14: Let W be the general solution of a homogeneous system $AX=0$, and suppose that the echelon form of the homogeneous system has s free variables. Let $u_1, u_2, \ldots, u_s$ be the solutions obtained by setting one of the free variables equal to 1 (or any nonzero constant) and the remaining free variables equal to 0. Then $\dim W = s$, and the vectors $u_1, u_2, \ldots, u_s$ form a basis of W.

We emphasize that the general solution W may have many bases, and that Theorem 3.12 only gives us one such basis.

EXAMPLE 3.18 Find the dimension and a basis for the general solution W of the homogeneous system

$$\begin{aligned} x_1+\ 2x_2-\ 3x_3+2x_4-\ 4x_5&=0\\ 2x_1+\ 4x_2-\ 5x_3+\ x_4-\ 6x_5&=0\\ 5x_1+10x_2-13x_3+4x_4-16x_5&=0 \end{aligned}$$

First reduce the system to echelon form. Apply the following operations:

"Replace L_2 by $-2L_1 + L_2$" and "Replace L_3 by $-5L_1 + L_3$" and then "Replace L_3 by $-2L_2 + L_3$"

These operations yield

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 - 4x_5 &= 0 \\ x_3 - 3x_4 + 2x_5 &= 0 \\ 2x_3 - 6x_4 + 4x_5 &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 - 4x_5 &= 0 \\ x_3 - 3x_4 + 2x_5 &= 0 \end{aligned}$$

The system in echelon form has three free variables, x_2, x_4, x_5; hence, $\dim W = 3$. Three solution vectors that form a basis for W are obtained as follows:

(1) Set $x_2 = 1$, $x_4 = 0$, $x_5 = 0$. Back-substitution yields the solution $u_1 = (-2, 1, 0, 0, 0)$.
(2) Set $x_2 = 0$, $x_4 = 1$, $x_5 = 0$. Back-substitution yields the solution $u_2 = (7, 0, 3, 1, 0)$.
(3) Set $x_2 = 0$, $x_4 = 0$, $x_5 = 1$. Back-substitution yields the solution $u_3 = (-2, 0, -2, 0, 1)$.

The vectors $u_1 = (-2, 1, 0, 0, 0)$, $u_2 = (7, 0, 3, 1, 0)$, $u_3 = (-2, 0, -2, 0, 1)$ form a basis for W.

Remark: Any solution of the system in Example 3.18 can be written in the form

$$\begin{aligned} au_1 + bu_2 + cu_3 &= a(-2, 1, 0, 0, 0) + b(7, 0, 3, 1, 0) + c(-2, 0, -2, 0, 1) \\ &= (-2a + 7b - 2c, \quad a, \quad 3b - 2c, \quad b, \quad c) \end{aligned}$$

or

$$x_1 = -2a + 7b - 2c, \qquad x_2 = a, \qquad x_3 = 3b - 2c, \qquad x_4 = b, \qquad x_5 = c$$

where a, b, c are arbitrary constants. Observe that this representation is nothing more than the parametric form of the general solution under the choice of parameters $x_2 = a$, $x_4 = b$, $x_5 = c$.

Nonhomogeneous and Associated Homogeneous Systems

Let $AX = B$ be a nonhomogeneous system of linear equations. Then $AX = 0$ is called the *associated homogeneous system*. For example,

$$\begin{aligned} x + 2y - 4z &= 7 \\ 3x - 5y + 6z &= 8 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x + 2y - 4z &= 0 \\ 3x - 5y + 6z &= 0 \end{aligned}$$

show a nonhomogeneous system and its associated homogeneous system.

The relationship between the solution U of a nonhomogeneous system $AX = B$ and the solution W of its associated homogeneous system $AX = 0$ is contained in the following theorem.

THEOREM 3.15: Let v_0 be a particular solution of $AX = B$ and let W be the general solution of $AX = 0$. Then the following is the general solution of $AX = B$:

$$U = v_0 + W = \{v_0 + w : w \in W\}$$

That is, $U = v_0 + W$ is obtained by adding v_0 to each element in W. We note that this theorem has a geometrical interpretation in $\mathbf{R}^3$. Specifically, suppose W is a line through the origin O. Then, as pictured in Fig. 3-4, $U = v_0 + W$ is the line parallel to W obtained by adding v_0 to each element of W. Similarly, whenever W is a plane through the origin O, then $U = v_0 + W$ is a plane parallel to W.