7. Let a, b, c be positive integers. Prove that there is no solution of $ax + by = c$ in positive integers if $a + b > c$.

8. If $ax + by = c$ is solvable, prove that it has a solution x_0, y_0 with $0 \leqslant x_0 < |b|$.

9. Prove that $ax + by = a + c$ is solvable if and only if $ax + by = c$ is solvable.

10. Prove that $ax + by = c$ is solvable if and only if $(a, b) = (a, b, c)$.

11. Given that $ax + by = c$ has two solutions, (x_0, y_0) and (x_1, y_1) with $x_1 = 1 + x_0$, and given that $(a, b) = 1$, prove that $b = \pm 1$.

12. A positive integer is called *powerful* if $p^2 | a$ whenever $p | a$. Show that a is powerful if and only if a can be expressed in the form $a = b^2c^3$ where b and c are positive integers.

13. Let a, b, c be positive integers such that $g | c$, where $g = \text{g.c.d.}(a, b)$, and let N denote the number of solutions of (5.1) in non-negative integers. Show that $N = [y_1 g/a] + [x_1 g/b] + 1 = gc/(ab) + 1 - \{y_1 g/a\} - \{x_1 g/b\}$.

14. Let a, b, c be positive integers. Assuming that $g | c$ and that $cg/(ab)$ is an integer, prove that $N = 1 + cg/(ab)$, and that $P = -1 + cg/(ab)$.

15. Let a, b, c be positive integers. Assuming that $g | c$ but that $cg/(ab)$ is not an integer, prove that $P = [cg/(ab)]$ or $P = [cg/(ab)] + 1$, and that $N = [cg/(ab)]$ or $N = [cg/(ab)] + 1$. Assuming further that $a | c$, show that $N = [cg/(ab)] + 1$ and that $P = [cg/(ab)]$. (H)

*16. Let a and b be positive integers with $\text{g.c.d.}(a, b) = 1$. Let $\mathscr{S}$ denote the set of all integers that may be expressed in the form $ax + by$ where x and y are non-negative integers. Show that $c = ab - a - b$ is not a member of $\mathscr{S}$, but that every integer larger than c is a member of $\mathscr{S}$.

*17. Find necessary and sufficient conditions that

$$x + b_1 y + c_1 z = d_1, \qquad x + b_2 y + c_2 z = d_2$$

have at least one simultaneous solution in integers x, y, z, assuming that the coefficients are integers with $b_1 \neq b_2$.

5.2 SIMULTANEOUS LINEAR EQUATIONS

Let $a_1, a_2, \cdots, a_n$ be integers, not all 0, and suppose we wish to find all solutions in integers of the equation

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = c.$$

As in Theorem 5.1, we may show that such solutions exist if and only if g.c.d.$(a_1, a_2, \cdots, a_n)$ divides c. The numerical technique exposed in the preceding section also extends easily to larger values of n.

Example 3 Find all solutions in integers of $2x + 3y + 4z = 5$.

Solution We write

$$\begin{array}{cccc} 2 & 3 & 4 & 5 \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \end{array} \quad \rightarrow \quad \begin{array}{cccc} 2 & 1 & 0 & 5 \\ 1 & -1 & -2 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \end{array} \quad \rightarrow \quad \begin{array}{cccc} 0 & 1 & 0 & 5 \\ 3 & -1 & -2 & \\ -2 & 1 & 0 & \\ 0 & 0 & 1 & \end{array}$$

This last array represents simultaneous equations involving three new variables, say t, u, v. The first line gives the condition $u = 5$. On substituting this in the lower lines, we find that every solution of the given equation in integers may be expressed in the form

$$\begin{aligned} x &= 3t - 2v - 5 \\ y &= -2t \qquad + 5 \\ z &= \qquad v \end{aligned}$$

where t and v are integers. From the nature of the changes of variables made, we know that triples (x, y, z) of integers satisfying the given equation are in one-to-one correspondence with triples of integers (t, u, v) for which $u = 5$. Hence each solution of the given equation in integers is given by a unique pair of integers (t, v).

We now consider the problem of treating simultaneous equations. Suppose we have two equations, say

$$\begin{aligned} A &= B, \\ C &= D. \end{aligned} \tag{5.8}$$

By multiplying the first equation by m and adding the result to the second equation, we may obtain a new pair of equations,

$$\begin{aligned} A &= B, \\ C + mA &= D + mB. \end{aligned} \tag{5.9}$$

This pair of equations is equivalent to the original pair (5.8). Here m may be any real number, but since our interest is in equations with integral coefficients, we shall restrict m to be an integer. Similarly, the equation $A = B$ is equivalent to $cA = cB$ provided that $c \neq 0$. Again, since our interest is in equations with integral coefficients, we restrict c to the values $c = \pm 1$. Finally, we may rearrange a collection of equations without altering their significance. Hence we have at our disposal three row operations which we may apply to a system of equations:

(R1) Add an integral multiple m of one equation to another;

(R2) Exchange two equations;

(R3) Multiply both sides of an equation by -1.

By applying these operations in conjunction with the column operations considered in the preceding section, we may determine the integral solutions of a system of linear equations.

Example 4 Find all solutions in integers of the simultaneous equations

$$20x + 44y + 50z = 10,$$

$$17x + 13y + 11z = 19.$$

Solution Among the coefficients of x, y, and z, the coefficient 11 is smallest. Using operation (C1) and the division algorithm (rounding to the nearest integer), reduce the coefficients of x and y in the second row (mod 11):

$$\begin{array}{rrrr} 20 & 44 & 50 & 10 \\ 17 & 13 & 11 & 19 \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \end{array} \quad \rightarrow \quad \begin{array}{rrrr} -80 & -6 & 50 & 10 \\ -5 & 2 & 11 & 19 \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ -2 & -1 & 1 & \end{array}$$

The coefficient of least absolute value is now in the second row and second column. We use operation (C1) to reduce the other coefficients in the second row (mod 2):

$$\rightarrow \quad \begin{array}{rrrr} -98 & -6 & 80 & 10 \\ 1 & 2 & 1 & 19 \\ 1 & 0 & 0 & \\ 3 & 1 & -5 & \\ -5 & -1 & 6 & \end{array}$$

There are now two coefficients of minimal absolute value. We use the one in the first column as our pivot and use operation (C1) to reduce the other coefficients in the second row:

$$\rightarrow \quad \begin{array}{rrrr} -98 & 190 & 178 & 10 \\ 1 & 0 & 0 & 19 \\ 1 & -2 & -1 & \\ 3 & -5 & -8 & \\ -5 & 9 & 11 & \end{array}$$

The coefficient of least nonzero absolute value is unchanged, so we switch to operation (R1) to reduce the coefficient $-98 \pmod 1$, and then we use (R2) to interchange the two rows:

$$\rightarrow \quad \begin{array}{rrrr} 0 & 190 & 178 & 1872 \\ 1 & 0 & 0 & 19 \\ 1 & -2 & -1 & \\ 3 & -5 & -8 & \\ -5 & 9 & 11 & \end{array} \qquad \rightarrow \quad \begin{array}{rrrr} 1 & 0 & 0 & 19 \\ 0 & 190 & 178 & 1872 \\ 1 & -2 & -1 & \\ 3 & -5 & -8 & \\ -5 & 9 & 11 & \end{array}$$

We now ignore the first row and first column. Among the remaining coefficients, the one of least nonzero absolute value is 178. We use operation (C1) to reduce $190 \pmod{178}$, obtaining a remainder 12. Then we use (C1) to reduce $178 \pmod{12}$, obtaining a remainder -2:

$$\rightarrow \quad \begin{array}{rrrr} 1 & 0 & 0 & 19 \\ 0 & 12 & 178 & 1872 \\ 1 & -1 & -1 & \\ 3 & 3 & -8 & \\ -5 & -2 & 11 & \end{array} \qquad \rightarrow \quad \begin{array}{rrrr} 1 & 0 & 0 & 19 \\ 0 & 12 & -2 & 1872 \\ 1 & -1 & 14 & \\ 3 & 3 & -53 & \\ -5 & -2 & 41 & \end{array}$$

Next we use (C2) to reduce $12 \pmod 2$. Then we use (C2) to interchange the second and third columns, and finally use (C3) to replace -2 by 2:

$$\rightarrow \quad \begin{array}{rrrr} 1 & 0 & 0 & 19 \\ 0 & 0 & -2 & 1872 \\ 1 & 83 & 14 & \\ 3 & -315 & -53 & \\ -5 & 244 & 41 & \end{array} \qquad \rightarrow \quad \begin{array}{rrrr} 1 & 0 & 0 & 19 \\ 0 & 2 & 0 & 1872 \\ 1 & -14 & 83 & \\ 3 & 53 & -315 & \\ -5 & -41 & 244 & \end{array}$$

Let the variables in our new set of equations be called t, u, and v. The two original equations have been replaced by the two new equations $1 \cdot t = 19$ and $2 \cdot u = 1872$. This fixes the values of t and u. Since $1 | 19$ and $2 | 1872$, these values are integers: $t = 19$, $u = 936$. With these values for t and u, the bottom three rows above give the equations

$$\begin{aligned} x &= t - 14u + 83v = 83v - 13085, \\ y &= 3t + 53u - 315v = -315v + 49665, \\ z &= -5t - 41u + 244v = 244v - 38471. \end{aligned}$$

By making the further change of variable $w = v - 158$ we may adjust the constant terms, so that

$$\begin{aligned} x &= 83w + 29, \\ y &= -315w - 105, \\ z &= 244w + 81. \end{aligned}$$

As integral solutions of the given equations are in one-to-one correspondence with integral values of w, we have achieved our goal.

To demonstrate that this procedure will succeed in general, we describe the strategy more precisely. Suppose we wish to parameterize all integral solutions of a family of m linear equations in n variables,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{5.10}$$

We assume that the a_{ij} and the b_i are integers, with not all $a_{ij} = 0$. Our object is to find an equivalent family of m equations in n equivalent variables that is diagonal, in the sense that the new coefficients a_{ij} vanish whenever $i \neq j$. Let $A = [a_{ij}]$ be the $m \times n$ matrix of given coefficients, let $X = [x_j]$ denote the $n \times 1$ matrix (or column vector) of variables, and let $B = [b_i]$ be the $m \times 1$ matrix (or column vector) of given constant terms. Then the given equations may be expressed as the single matrix equation $AX = B$. If we let $V = [v_{ij}]$ be the $n \times n$ matrix that expresses our original variables in terms of our new variables $Y = [y_j]$, then $VY = X$. Initially, $V = I$, the identity matrix. We describe a reduction step that transforms A into a matrix $A' = [a'_{ij}]$ with the property that $a'_{11} \geqslant 0$, $a'_{1j} = 0$ for $j > 1$, and $a'_{i1} = 0$ for $i > 1$. By repeated use of this reduction step, A is eventually transformed into a diagonal matrix whose diagonal entries are non-negative. As we perform row and column operations on A, we obtain a sequence of coefficient matrices. Let μ denote the minimal absolute value of non-zero elements of the current coefficient matrix. Locating an element of absolute value μ, say in position (i_0, j_0), we use operation (C1) or operation (R1) to reduce the other coefficients in row i_0 or column j_0. This gives rise to a new coefficient matrix with a strictly smaller value of μ, unless all the other coefficients in row i_0 and column

j_0 are 0. Since μ can take on only positive integral values, this latter situation must eventually arise. Then we use operations (R2) and (C2) to move the coefficient from location (i_0, j_0) to $(1,1)$. If the coefficient is negative, we use (C3) to reverse the sign. Whenever we apply a column operation to the coefficient matrix A, we also apply the same column operation to V, and whenever we apply a row operation to A, we apply the same row operation to B. The reduction procedure will terminate prematurely if in the submatrix that remains to be treated all elements are 0. Thus we obtain a diagonal matrix with positive entries in the first r rows, and 0's elsewhere. In developing standard linear algebra over $\mathbb{R}$ it is found that the rank of a matrix is invariant under row or column operations. Since the row and column operations we are using here are a proper subset of those used in linear algebra over $\mathbb{R}$, the rank is invariant in the present situation, as well. As the rank of a diagonal matrix is simply equal to the number of nonzero elements, we see that the number r is the rank of the matrix A given originally.

Caution At all stages of the reduction process, the column operations must involve only columns 1 through n. Similarly, the row operations must involve only rows 1 through m.

In summary, the change of variables $VY = X$ has the property that n-tuples X of integers are in one-to-one correspondence with n-tuples Y of integers. The m conditions (5.10) on the variables x_j are equivalent to the m conditions

$$d_j y_j = b_j' \qquad (1 \leqslant j \leqslant r), \tag{5.11}$$

$$b_j' = 0 \qquad (r < j \leqslant m). \tag{5.12}$$

Here the d_j are the diagonal entries of the new coefficient matrix, and the b_j' are the new constant terms. In order that integral solutions should exist, it is necessary and sufficient that (5.12) holds, and that

$$d_j | b_j' \qquad (1 \leqslant j \leqslant r). \tag{5.13}$$

If (5.12) holds but (5.13) fails for some $j \leqslant r$, then there exist rational solutions but no integral solution. If (5.12) fails for some $j > r$ then the original equations are inconsistent, and then (5.10) has no solution in real variables. If (5.11) and (5.12) hold and $r = n$, then the integral solution is unique (and indeed this is the unique real solution). If (5.12) and (5.13) hold but $r < n$ then there are infinitely many integral solutions, parameterized by the free integral variables $y_{r+1}, y_{r+2}, \cdots, y_n$.

As we experienced in Example 4, the coefficients encountered during the reduction process may be much larger than the coefficients originally given. (It is not known precisely how much larger, but it is believed that they may be *very much* larger. It is interesting to consider how the reduction process might be modified in order to minimize this phenomenon.) However, this problem does not arise when the method is applied to systems of simultaneous congruences (mod q) instead of simultaneous equations, for then coefficients may be reduced (mod q) during the reduction process. Here q may be any integer > 1, but it is imperative that each congruence involves the same modulus q.

Example 5 Find all solutions of the simultaneous congruences

$$\begin{aligned} 3x \phantom{{}-y} + 3z &\equiv 1 \pmod{5}, \\ 4x - y + z &\equiv 3 \pmod{5}. \end{aligned}$$

Solution We construct an array of coefficients as before. Using operation (C1), we add the third column to both columns 1 and 2.

$$\begin{array}{rrrr} 3 & 0 & 3 & 1 \\ 4 & -1 & 1 & 3 \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \end{array} \quad \rightarrow \quad \begin{array}{rrrr} 1 & 3 & 3 & 1 \\ 0 & 0 & 1 & 3 \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ 1 & 1 & 1 & \end{array}$$

Using (R1), we multiply the second row by 2 and add the result to the first row. Then we interchange the first and third columns and the first and second rows.

$$\rightarrow \quad \begin{array}{rrrr} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ 1 & 1 & 1 & \end{array} \quad \rightarrow \quad \begin{array}{rrrr} 1 & 0 & 0 & 3 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 1 & \\ 0 & 1 & 0 & \\ 1 & 1 & 1 & \end{array}$$

Next we multiply the third column by 2 and add the result to the second column, and then interchange the second and third columns.

$$\rightarrow \quad \begin{array}{rrrr} 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 2 & 1 & \\ 0 & 1 & 0 & \\ 1 & 3 & 1 & \end{array} \quad \rightarrow \quad \begin{array}{rrrr} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 2 & \\ 0 & 0 & 1 & \\ 1 & 1 & 3 & \end{array}$$

Thus we arrive at a new system of congruences, in variables t, u, v, say. We see that $t \equiv 3 \pmod 5$, $u \equiv 2 \pmod 5$, while v can take any value (mod 5). Thus the given system has five solutions, given by

$$\begin{aligned} x &\equiv \quad u + 2v \equiv 2v + 2 \pmod 5, \\ y &\equiv \qquad\quad v \equiv \ v \qquad \pmod 5, \\ z &\equiv t + u + 3v \equiv 3v \qquad \pmod 5. \end{aligned}$$

In general, the system of simultaneous congruences

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\equiv b_1 \pmod q, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &\equiv b_2 \pmod q, \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad &\quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\equiv b_m \pmod q, \end{aligned} \tag{5.14}$$

has a solution $(\bmod\ q)$ if and only if

$$\text{g.c.d.}\,(d_j, q) | b_j' \qquad (1 \leqslant j \leqslant r), \tag{5.15}$$

$$b_j' \equiv 0 \pmod q \qquad (r < j \leqslant m). \tag{5.16}$$

Note that these conditions may hold while (5.12) fails. In such a case the congruences (5.14) have a simultaneous solution even though the equations (5.10) have no real solution. On the other hand, if (5.10) has a real solution then (5.12) holds. If we take q to be a multiple of all of the d_j then the conditions (5.15) are equivalent to (5.13). This gives the following important result.

Theorem 5.2 *If the system of linear equations* (5.10) *has a real solution, and if the system of congruences* (5.14) *has a solution for every modulus* q, *then the equations* (5.10) *have an integral solution.*

We have actually proved more, since we can determine a particular q that suffices. (For a more precise characterization of this q in terms of the original coefficients, see Problem 11 at the end of this section.) The converse of the theorem is obvious, for if a system of equations (even nonlinear equations) has an integral solution then this solution is both a real solution and also a congruential solution for any q. We speak of the congruential and real solutions as "local," while an integral solution is "global." In this parlance, Theorem 5.2 may be expressed by saying that the equations (5.10) have a global solution if they are everywhere locally solvable.

While our main aims in this Section have been achieved, further insights may be gained by making greater use of linear algebra. Suppose that a particular row operation, applied to the $m \times n$ matrix A, gives the matrix A'. Let R denote the matrix obtained by applying this same row operation to the $m \times m$ identity matrix I_m. Then $A' = RA$. We call such a matrix R an *elementary row matrix*. Note that the elementary row matrices here form a proper subset of the elementary row matrices defined in standard linear algebra over $\mathbb{R}$, since we have restricted the row operations that are allowed. Similarly, if a particular column operation takes A to A'' and I_n to C, then $A'' = AC$, and we call C an *elementary column matrix*. Thus the sequence of row and column operations that we have performed in our reduction process may be expressed by matrix multiplication,

$$R_g R_{g-1} \cdots R_2 R_1 A C_1 C_2 \cdots C_{h-1} C_h = D, \tag{5.17}$$

where D is an $m \times n$ diagonal matrix. (Note that a diagonal matrix is not necessarily square.) The matrix V that allows us to express the original variables X in terms of our new variables Y is constructed by applying the same column operations to the identity matrix. That is,

$$V = C_1 C_2 \cdots C_{h-1} C_h. \tag{5.18}$$

Similarly, the new constant terms B' obtained at the end of the reduction process are created by applying the row operations to the original set B of constant terms, so that

$$B' = R_g R_{g-1} \cdots R_2 R_1 B. \tag{5.19}$$

It is useful to characterize those matrices that may be written as products of our elementary row or column matrices.

Definition 5.1 *A square matrix U with integral elements is called* unimodular *if* $\det(U) = \pm 1$.

Theorem 5.3 *Let U be an $m \times m$ matrix with integral elements. Then the following are equivalent:*

(i) U is unimodular;

(ii) The inverse matrix U^{-1} exists and has integral elements;

(iii) U may be expressed as a product of elementary row matrices.

$$U = R_g R_{g-1} \cdots R_2 R_1;$$

(iv) U may be expressed as a product of elementary column matrices,

$$U = C_1 C_2 \cdots C_{h-1} C_h.$$

If U and V are $m \times m$ unimodular matrices, then so also is UV, in view of (3.6). Moreover, U^{-1} is unimodular, by (ii) above. Thus the collection of all $m \times m$ unimodular matrices forms a group.

Proof We first show that (i) implies (ii). From the definition of the adjoint matrix U^{adj} it is evident that if U has integral elements then so does U^{adj}. Since $U^{-1} = U^{\text{adj}}/\det(U)$, it follows that U^{-1} has integral elements if $\det(U) = \pm 1$. We next show that (ii) implies (i). Since $UU^{-1} = I$, it follows by (3.6) that $\det(U)\det(U^{-1}) = \det(I) = 1$. But $\det(U)$ is an integer if U has integral elements, so from (ii) we deduce that both $\det(U)$ and $\det(U^{-1})$ are integers. That is, $\det(U)$ divides 1. As the only divisors of 1 are ± 1, it follows that U is unimodular. Next we show that (iii) implies (i). It is easy to verify that an elementary row matrix is unimodular. From (3.6) it is evident that the product of two unimodular matrices is again unimodular. Thus if $U = R_g R_{g-1} \cdots R_2 R_1$, then U is unimodular.

To show that (i) implies (iii), we first show that if A is an $m \times n$ matrix with integral elements then there exist elementary row matrices such that

$$A = R_1 R_2 \cdots R_{g-1} R_g T \tag{5.20}$$

where T is an upper-triangular $m \times n$ matrix with integral elements. We proceed as in Gaussian elimination in elementary linear algebra, except that we restrict ourselves to the row operations (R1), (R2), and (R3). We apply these row operations to A as follows. In the first column containing nonzero elements, say the first column, we apply the division algorithm and (R1) until only one element in this column is nonzero. By means of (R2) we may place this nonzero entry in the first row. By (R3) we may arrange that this element is positive. We now repeat this process on the columns to the right of the one just considered, but we ignore the first row. Thus the second column operated on may have two nonzero elements, in the first and second rows. Continuing in this manner, we arrive at an upper triangular matrix T. That is, $T = R_g R_{g-1} \cdots R_2 R_1 A$ for suitable elementary row matrices R_i. Hence $A = R_1^{-1} R_2^{-2} \cdots R_{g-1}^{-1} R_g^{-1} T$. Since the inverse of an elementary row matrix is again an elementary row matrix, we have now expressed A in the desired form (5.20).

To complete the proof that (i) implies (iii), we take $A = U$ in (5.20). Applying (3.6), we deduce that $\det(T) = \pm 1$. But since T is upper-triangular, $\det(T)$ is the product of its diagonal elements. As these diagonal elements are non-negative integers, it follows that each diagonal element is 1. With this established, we may now apply the row operation (R1) to T to clear all entries above the diagonal, leaving us with the identity matrix

I_m. That is, T is the product of elementary row matrices, and hence by (5.20), so also is U.

The equivalence of (i) and (iv) may be established similarly. Alternatively, we observe that R is an elementary row matrix if and only if R^t is an elementary column matrix. (Here R^t denotes the transpose of R.) If U is unimodular then U^t is unimodular, and by (iii) we deduce that $U^t = R_g R_{g-1} \cdots R_2 R_1$ for suitable elementary row matrices R_i. Hence $U = R_1^t R_2^t \cdots R_{g-1}^t R_g^t$, a product of column matrices.

We call two $m \times n$ matrices A and A' *equivalent*, and write $A \sim A'$, if there exists an $m \times m$ unimodular matrix U and an $n \times n$ unimodular matrix V such that $A' = UAV$. This is an equivalence relation in the usual sense. With this machinery in hand, we may express (5.17) more compactly by saying that any matrix A is equivalent to a diagonal matrix, say $UAV = D$. Then $A = U^{-1}DV^{-1}$. Writing (5.10) in the form $AX = B$, we deduce that $U^{-1}DV^{-1}X = B$. On putting $Y = V^{-1}X$, $UB = B'$, we are led immediately to the conclusion that (5.10) is equivalent to $DY = B'$, which is precisely the content of (5.11) and (5.12).

Owing to ambiguities in our reduction process, the diagonal matrix D that we have found to be equivalent to A is not uniquely defined. Moreover, two different diagonal matrices may be equivalent, as we see from the example

$$\begin{bmatrix} 1 & 1 \\ -3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} -1 & -3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}.$$

However, it is known that among the diagonal matrices equivalent to a given matrix A there is a unique one whose nonzero elements $s_1, s_2, \cdots, s_r$ are positive and satisfy the divisibility relations $s_1 | s_2, s_2 | s_3, \cdots, s_{r-1} | s_r$. This diagonal matrix S is the *Smith normal form of* A, named for the nineteenth-century English mathematician H. J. S. Smith. The numbers s_i, $1 \leqslant i \leqslant r$, are called the *invariant factors* of A. A proof that every $m \times n$ matrix A is equivalent to a unique matrix S in Smith normal form is outlined in Problems 4–9.

PROBLEMS

1. Find all solutions in integers of the system of equations

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

2. For what integers a, b, and c does the system of equations

$$\begin{aligned} x_1 + 2x_2 + \ 3x_3 + \ 4x_4 &= a, \\ x_1 + 4x_2 + \ 9x_3 + 16x_4 &= b, \\ x_1 + 8x_2 + 27x_3 + 64x_4 &= c \end{aligned}$$

have a solution in integers? What are the solutions if $a = b = c = 1$?

3. Suppose that the system of congruences (5.14) has a solution. Show that if q is prime then the number of solutions is a power of q.

*4. Let a and b be positive integers, and put $g =$ g.c.d. (a, b), $h =$ l.c.m. (a, b). Show that $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \sim \begin{bmatrix} g & 0 \\ 0 & h \end{bmatrix}$.

*5. Using the preceding problem, or otherwise, show that if D is a diagonal matrix with integral elements then there is a diagonal matrix S in Smith normal form such that $D \sim S$. Deduce that every $m \times n$ matrix A with integral elements is equivalent to a matrix S in Smith normal form.

*6. Let A be an $m \times n$ matrix with integral elements, and let r denote the rank of A. For $1 \leqslant k \leqslant r$, let $d_k(A)$ be the greatest common divisor of the determinants of all $k \times k$ minors of A. The numbers $d_k(A)$ are called the *determinantal divisors* of A. Let R be an elementary unimodular row matrix, and put $A' = RA$. Show that A and A' have the same determinantal divisors.

*7. Use the preceding problem to show that if A and B are equivalent matrices then they have the same determinantal divisors.

*8. Let S be a matrix in Smith normal form whose positive diagonal elements are $s_1, s_2, \cdots, s_r$. Show that $d_1(S) = s_1$, $d_2(S) = s_1 s_2, \cdots, d_r(S) = s_1 s_2 \cdots s_r$. For convenience, put $d_0(S) = 1$. Deduce that $s_k = d_k(S)/d_{k-1}(S)$ for $1 \leqslant k \leqslant r$.

*9. Let S and S' be two $m \times n$ matrices in Smith normal form. Using the preceding problems, show that if $S \sim S'$ then $S = S'$. Conclude that the Smith normal form of an $m \times n$ matrix A is unique.

*10. Show that if two $m \times n$ matrices A and A' have the same rank and the same determinantal divisors then $A \sim A'$.

*11. Suppose that the system of equations (5.10) has real solutions, and that the system of congruences (5.14) has a solution when $q = d_r(A)/d_{r-1}(A)$. Show that the equations (5.10) have an integral solution. Show also that this is the least integer q for which this conclusion may be drawn.

*12. Let A be an $n \times n$ matrix with integral elements and nonzero determinant. Then the elements of A^{-1} are rational numbers. Show that the least common denominator of these elements is $d_n(A)/d_{n-1}(A)$.