1.3. Matrices and Elementary Row Operations

One cannot fail to notice that in forming linear combinations of linear equations there is no need to continue writing the 'unknowns' $x_1, \ldots, x_n$, since one actually computes only with the coefficients A_{ij} and the scalars y_i. We shall now abbreviate the system (1-1) by

$$AX = Y$$

where

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}$$

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

We call A the **matrix of coefficients** of the system. Strictly speaking, the rectangular array displayed above is not a matrix, but is a representation of a matrix. An $m \times n$ **matrix over the field** F is a function A from the set of pairs of integers (i, j), $1 \leq i \leq m$, $1 \leq j \leq n$, into the field F. The **entries** of the matrix A are the scalars $A(i, j) = A_{ij}$, and quite often it is most convenient to describe the matrix by displaying its entries in a rectangular array having m rows and n columns, as above. Thus X (above) is, or defines, an $n \times 1$ matrix and Y is an $m \times 1$ matrix. For the time being, $AX = Y$ is nothing more than a shorthand notation for our system of linear equations. Later, when we have defined a multiplication for matrices, it will mean that Y is the product of A and X.

We wish now to consider operations on the rows of the matrix A which correspond to forming linear combinations of the equations in the system $AX = Y$. We restrict our attention to three **elementary row operations** on an $m \times n$ matrix A over the field F:

1. multiplication of one row of A by a non-zero scalar c;
2. replacement of the rth row of A by row r plus c times row s, c any scalar and $r \neq s$;
3. interchange of two rows of A.

An elementary row operation is thus a special type of function (rule) e which associated with each $m \times n$ matrix A an $m \times n$ matrix $e(A)$. One can precisely describe e in the three cases as follows:

1. $e(A)_{ij} = A_{ij}$ if $i \neq r$, $e(A)_{rj} = cA_{rj}$.
2. $e(A)_{ij} = A_{ij}$ if $i \neq r$, $e(A)_{rj} = A_{rj} + cA_{sj}$.
3. $e(A)_{ij} = A_{ij}$ if i is different from both r and s, $e(A)_{rj} = A_{sj}$, $e(A)_{sj} = A_{rj}$.

In defining $e(A)$, it is not really important how many columns A has, but the number of rows of A is crucial. For example, one must worry a little to decide what is meant by interchanging rows 5 and 6 of a 5×5 matrix. To avoid any such complications, we shall agree that an elementary row operation e is defined on the class of all $m \times n$ matrices over F, for some fixed m but any n. In other words, a particular e is defined on the class of all m-rowed matrices over F.

One reason that we restrict ourselves to these three simple types of row operations is that, having performed such an operation e on a matrix A, we can recapture A by performing a similar operation on $e(A)$.

Theorem 2. *To each elementary row operation* e *there corresponds an elementary row operation* e_1, *of the same type as* e, *such that* $e_1(e(A)) = e(e_1(A)) = A$ *for each* A. *In other words, the inverse operation (function) of an elementary row operation exists and is an elementary row operation of the same type.*

Proof. (1) Suppose e is the operation which multiplies the rth row of a matrix by the non-zero scalar c. Let e_1 be the operation which multiplies row r by c^{-1}. (2) Suppose e is the operation which replaces row r by row r plus c times row s, $r \neq s$. Let e_1 be the operation which replaces row r by row r plus $(-c)$ times row s. (3) If e interchanges rows r and s, let $e_1 = e$. In each of these three cases we clearly have $e_1(e(A)) = e(e_1(A)) = A$ for each A. ∎

Definition. *If* A *and* B *are* $m \times n$ *matrices over the field* F, *we say that* B **is row-equivalent to** A *if* B *can be obtained from* A *by a finite sequence of elementary row operations.*

Using Theorem 2, the reader should find it easy to verify the following. Each matrix is row-equivalent to itself; if B is row-equivalent to A, then A is row-equivalent to B; if B is row-equivalent to A and C is row-equivalent to B, then C is row-equivalent to A. In other words, row-equivalence is an equivalence relation (see Appendix).

Theorem 3. *If* A *and* B *are row-equivalent* $m \times n$ *matrices, the homogeneous systems of linear equations* $AX = 0$ *and* $BX = 0$ *have exactly the same solutions.*

Proof. Suppose we pass from A to B by a finite sequence of elementary row operations:

$$A = A_0 \to A_1 \to \cdots \to A_k = B.$$

It is enough to prove that the systems $A_jX = 0$ and $A_{j+1}X = 0$ have the same solutions, i.e., that one elementary row operation does not disturb the set of solutions.

So suppose that B is obtained from A by a single elementary row operation. No matter which of the three types the operation is, (1), (2), or (3), each equation in the system $BX = 0$ will be a linear combination of the equations in the system $AX = 0$. Since the inverse of an elementary row operation is an elementary row operation, each equation in $AX = 0$ will also be a linear combination of the equations in $BX = 0$. Hence these two systems are equivalent, and by Theorem 1 they have the same solutions. ∎

Example 5. Suppose F is the field of rational numbers, and

$$A = \begin{bmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix}.$$

We shall perform a finite sequence of elementary row operations on A, indicating by numbers in parentheses the type of operation performed.

$$\begin{bmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 0 & -2 & -1 & 7 \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 4 & 0 & -1 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 0 & -9 & 3 & 4 \\ 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 0 & \frac{15}{2} & -\frac{55}{2} \\ 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(1)}$$

$$\begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{(2)}$$

$$\begin{bmatrix} 0 & 0 & 1 & -\frac{11}{3} \\ 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{5}{3} \end{bmatrix}$$

The row-equivalence of A with the final matrix in the above sequence tells us in particular that the solutions of

$$\begin{aligned} 2x_1 - x_2 + 3x_3 + 2x_4 &= 0 \\ x_1 + 4x_2 \qquad\quad - x_4 &= 0 \\ 2x_1 + 6x_2 - x_3 + 5x_4 &= 0 \end{aligned}$$

and

$$\begin{aligned} x_3 - \tfrac{11}{3}x_4 &= 0 \\ x_1 \qquad\quad + \tfrac{17}{3}x_4 &= 0 \\ x_2 \qquad - \tfrac{5}{3}x_4 &= 0 \end{aligned}$$

are exactly the same. In the second system it is apparent that if we assign any rational value c to x_4 we obtain a solution $(-\frac{17}{3}c, \frac{5}{3}, \frac{11}{3}c, c)$, and also that every solution is of this form.

EXAMPLE 6. Suppose F is the field of complex numbers and

$$A = \begin{bmatrix} -1 & i \\ -i & 3 \\ 1 & 2 \end{bmatrix}.$$

In performing row operations it is often convenient to combine several operations of type (2). With this in mind

$$\begin{bmatrix} -1 & i \\ -i & 3 \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 2+i \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(1)} \begin{bmatrix} 0 & 1 \\ 0 & 3+2i \\ 1 & 2 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Thus the system of equations

$$\begin{aligned} -x_1 + ix_2 &= 0 \\ -ix_1 + 3x_2 &= 0 \\ x_1 + 2x_2 &= 0 \end{aligned}$$

has only the trivial solution $x_1 = x_2 = 0$.

In Examples 5 and 6 we were obviously not performing row operations at random. Our choice of row operations was motivated by a desire to simplify the coefficient matrix in a manner analogous to 'eliminating unknowns' in the system of linear equations. Let us now make a formal definition of the type of matrix at which we were attempting to arrive.

Definition. *An* m $\times$ n *matrix* R *is called* **row-reduced** *if:*

(a) *the first non-zero entry in each non-zero row of* R *is equal to* 1;

(b) *each column of* R *which contains the leading non-zero entry of some row has all its other entries* 0.

EXAMPLE 7. One example of a row-reduced matrix is the $n \times n$ (square) **identity matrix** I. This is the $n \times n$ matrix defined by

$$I_{ij} = \delta_{ij} = \begin{cases} 1, & \text{if} \quad i = j \\ 0, & \text{if} \quad i \neq j. \end{cases}$$

This is the first of many occasions on which we shall use the **Kronecker delta** (δ).

In Examples 5 and 6, the final matrices in the sequences exhibited there are row-reduced matrices. Two examples of matrices which are *not* row-reduced are:

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

The second matrix fails to satisfy condition (a), because the leading non-zero entry of the first row is not 1. The first matrix does satisfy condition (a), but fails to satisfy condition (b) in column 3.

We shall now prove that we can pass from any given matrix to a row-reduced matrix, by means of a finite number of elementary row operations. In combination with Theorem 3, this will provide us with an effective tool for solving systems of linear equations.

Theorem 4. *Every* m $\times$ n *matrix over the field* F *is row-equivalent to a row-reduced matrix.*

Proof. Let A be an $m \times n$ matrix over F. If every entry in the first row of A is 0, then condition (a) is satisfied in so far as row 1 is concerned. If row 1 has a non-zero entry, let k be the smallest positive integer j for which $A_{1j} \neq 0$. Multiply row 1 by A_{1k}^{-1}, and then condition (a) is satisfied with regard to row 1. Now for each $i \geq 2$, add $(-A_{ik})$ times row 1 to row i. Now the leading non-zero entry of row 1 occurs in column k, that entry is 1, and every other entry in column k is 0.

Now consider the matrix which has resulted from above. If every entry in row 2 is 0, we do nothing to row 2. If some entry in row 2 is different from 0, we multiply row 2 by a scalar so that the leading non-zero entry is 1. In the event that row 1 had a leading non-zero entry in column k, this leading non-zero entry of row 2 cannot occur in column k; say it occurs in column $k_r \neq k$. By adding suitable multiples of row 2 to the various rows, we can arrange that all entries in column k' are 0, except the 1 in row 2. The important thing to notice is this: In carrying out these last operations, we will not change the entries of row 1 in columns $1, \ldots, k$; nor will we change any entry of column k. Of course, if row 1 was identically 0, the operations with row 2 will not affect row 1.

Working with one row at a time in the above manner, it is clear that in a finite number of steps we will arrive at a row-reduced matrix. ∎

Exercises

1. Find all solutions to the system of equations

$$\begin{aligned} (1 - i)x_1 - ix_2 &= 0 \\ 2x_1 + (1 - i)x_2 &= 0. \end{aligned}$$

2. If

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

find all solutions of $AX = 0$ by row-reducing A.

3. If

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

find all solutions of $AX = 2X$ and all solutions of $AX = 3X$. (The symbol cX denotes the matrix each entry of which is c times the corresponding entry of X.)

4. Find a row-reduced matrix which is row-equivalent to

$$A = \begin{bmatrix} i & -(1+i) & 0 \\ 1 & -2 & 1 \\ 1 & 2i & -1 \end{bmatrix}.$$

5. Prove that the following two matrices are *not* row-equivalent:

$$\begin{bmatrix} 2 & 0 & 0 \\ a & -1 & 0 \\ b & c & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 2 \\ -2 & 0 & -1 \\ 1 & 3 & 5 \end{bmatrix}.$$

6. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be a 2×2 matrix with complex entries. Suppose that A is row-reduced and also that $a + b + c + d = 0$. Prove that there are exactly three such matrices.

7. Prove that the interchange of two rows of a matrix can be accomplished by a finite sequence of elementary row operations of the other two types.

8. Consider the system of equations $AX = 0$ where

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is a 2×2 matrix over the field F. Prove the following.

(a) If every entry of A is 0, then every pair (x_1, x_2) is a solution of $AX = 0$.

(b) If $ad - bc \neq 0$, the system $AX = 0$ has only the trivial solution $x_1 = x_2 = 0$.

(c) If $ad - bc = 0$ and some entry of A is different from 0, then there is a solution (x_1^0, x_2^0) such that (x_1, x_2) is a solution if and only if there is a scalar y such that $x_1 = yx_1^0$, $x_2 = yx_2^0$.

1.4. Row-Reduced Echelon Matrices

Until now, our work with systems of linear equations was motivated by an attempt to find the solutions of such a system. In Section 1.3 we established a standardized technique for finding these solutions. We wish now to acquire some information which is slightly more theoretical, and for that purpose it is convenient to go a little beyond row-reduced matrices.

Definition. *An* $m \times n$ *matrix* R *is called a* **row-reduced echelon matrix** *if:*