equivalent to A. We wish to show that $R = I$. That amounts to showing that the last row of R is not (identically) 0. Let

$$E = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

If the system $RX = E$ can be solved for X, the last row of R cannot be 0. We know that $R = PA$, where P is invertible. Thus $RX = E$ if and only if $AX = P^{-1}E$. According to (iii), the latter system has a solution. ∎

Corollary. *A square matrix with either a left or right inverse is invertible.*

Proof. Let A be an $n \times n$ matrix. Suppose A has a left inverse, i.e., a matrix B such that $BA = I$. Then $AX = 0$ has only the trivial solution, because $X = IX = B(AX)$. Therefore A is invertible. On the other hand, suppose A has a right inverse, i.e., a matrix C such that $AC = I$. Then C has a left inverse and is therefore invertible. It then follows that $A = C^{-1}$ and so A is invertible with inverse C. ∎

Corollary. *Let $A = A_1A_2 \cdots A_k$, where $A_1 \ldots, A_k$ are $n \times n$ (square) matrices. Then A is invertible if and only if each A_j is invertible.*

Proof. We have already shown that the product of two invertible matrices is invertible. From this one sees easily that if each A_j is invertible then A is invertible.

Suppose now that A is invertible. We first prove that A_k is invertible. Suppose X is an $n \times 1$ matrix and $A_kX = 0$. Then $AX = (A_1 \cdots A_{k-1})A_kX = 0$. Since A is invertible we must have $X = 0$. The system of equations $A_kX = 0$ thus has no non-trivial solution, so A_k is invertible. But now $A_1 \cdots A_{k-1} = AA_k^{-1}$ is invertible. By the preceding argument, A_{k-1} is invertible. Continuing in this way, we conclude that each A_j is invertible. ∎

We should like to make one final comment about the solution of linear equations. Suppose A is an $m \times n$ matrix and we wish to solve the system of equations $AX = Y$. If R is a row-reduced echelon matrix which is row-equivalent to A, then $R = PA$ where P is an $m \times m$ invertible matrix. The solutions of the system $AX = Y$ are exactly the same as the solutions of the system $RX = PY$ $(= Z)$. In practice, it is not much more difficult to find the matrix P than it is to row-reduce A to R. For, suppose we form the augmented matrix A' of the system $AX = Y$, with arbitrary scalars $y_1, \ldots, y_m$ occurring in the last column. If we then perform on A' a sequence of elementary row operations which leads from A to R, it will become evident what the matrix P is. (The reader should refer to Example 9 where we essentially carried out this process.) In particular, if A is a square matrix, this process will make it clear whether or not A is invertible and if A is invertible what the inverse P is. Since we have already given the nucleus of one example of such a computation, we shall content ourselves with a 2×2 example.

EXAMPLE 15. Suppose F is the field of rational numbers and

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

Then

$$\begin{bmatrix} 2 & -1 & y_1 \\ 1 & 3 & y_2 \end{bmatrix} \xrightarrow{(3)} \begin{bmatrix} 1 & 3 & y_2 \\ 2 & -1 & y_1 \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 1 & 3 & y_2 \\ 0 & -7 & y_1 - 2y_2 \end{bmatrix} \xrightarrow{(1)}$$
$$\begin{bmatrix} 1 & 3 & y_2 \\ 0 & 1 & \frac{1}{7}(2y_2 - y_1) \end{bmatrix} \xrightarrow{(2)} \begin{bmatrix} 1 & 0 & \frac{1}{7}(y_2 + 3y_1) \\ 0 & 1 & \frac{1}{7}(2y_2 - y_1) \end{bmatrix}$$

from which it is clear that A is invertible and

$$A^{-1} = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}.$$

It may seem cumbersome to continue writing the arbitrary scalars $y_1, y_2, \ldots$ in the computation of inverses. Some people find it less awkward to carry along two sequences of matrices, one describing the reduction of A to the identity and the other recording the effect of the same sequence of operations starting from the identity. The reader may judge for himself which is a neater form of bookkeeping.

EXAMPLE 16. Let us find the inverse of

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

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{12} & \frac{1}{12} \\ 0 & \frac{1}{12} & \frac{4}{45} \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ -\frac{1}{3} & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{12} & \frac{1}{12} \\ 0 & 0 & \frac{1}{180} \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{6} & -1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ -6 & 12 & 0 \\ 30 & -180 & 180 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} -9 & 60 & -60 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}.$$

It must have occurred to the reader that we have carried on a lengthy discussion of the rows of matrices and have said little about the columns. We focused our attention on the rows because this seemed more natural from the point of view of linear equations. Since there is obviously nothing sacred about rows, the discussion in the last sections could have been carried on using columns rather than rows. If one defines an elementary column operation and column-equivalence in a manner analogous to that of elementary row operation and row-equivalence, it is clear that each $m \times n$ matrix will be column-equivalent to a 'column-reduced echelon' matrix. Also each elementary column operation will be of the form $A \to AE$, where E is an $n \times n$ elementary matrix—and so on.

Exercises

1. Let

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

Find a row-reduced echelon matrix R which is row-equivalent to A and an invertible 3×3 matrix P such that $R = PA$.

2. Do Exercise 1, but with

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

3. For each of the two matrices

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

use elementary row operations to discover whether it is invertible, and to find the inverse in case it is.

4. Let

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

For which X does there exist a scalar c such that $AX = cX$?

5. Discover whether

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

is invertible, and find A^{-1} if it exists.

6. Suppose A is a 2×1 matrix and that B is a 1×2 matrix. Prove that $C = AB$ is not invertible.

7. Let A be an $n \times n$ (square) matrix. Prove the following two statements:

(a) If A is invertible and $AB = 0$ for some $n \times n$ matrix B, then $B = 0$.

(b) If A is not invertible, then there exists an $n \times n$ matrix B such that $AB = 0$ but $B \neq 0$.

8. Let

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

Prove, using elementary row operations, that A is invertible if and only if $(ad - bc) \neq 0$.

9. An $n \times n$ matrix A is called **upper-triangular** if $A_{ij} = 0$ for $i > j$, that is, if every entry below the main diagonal is 0. Prove that an upper-triangular (square) matrix is invertible if and only if every entry on its main diagonal is different from 0.

10. Prove the following generalization of Exercise 6. If A is an $m \times n$ matrix, B is an $n \times m$ matrix and $n < m$, then AB is not invertible.

11. Let A be an $m \times n$ matrix. Show that by means of a finite number of elementary row and/or column operations one can pass from A to a matrix R which is both 'row-reduced echelon' and 'column-reduced echelon,' i.e., $R_{ij} = 0$ if $i \neq j$, $R_{ii} = 1$, $1 \leq i \leq r$, $R_{ii} = 0$ if $i > r$. Show that $R = PAQ$, where P is an invertible $m \times m$ matrix and Q is an invertible $n \times n$ matrix.

12. The result of Example 16 suggests that perhaps the matrix

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{bmatrix}$$

is invertible and A^{-1} has integer entries. Can you prove that?

2. Vector Spaces

2.1. Vector Spaces

In various parts of mathematics, one is confronted with a set, such that it is both meaningful and interesting to deal with 'linear combinations' of the objects in that set. For example, in our study of linear equations we found it quite natural to consider linear combinations of the rows of a matrix. It is likely that the reader has studied calculus and has dealt there with linear combinations of functions; certainly this is so if he has studied differential equations. Perhaps the reader has had some experience with vectors in three-dimensional Euclidean space, and in particular, with linear combinations of such vectors.

Loosely speaking, linear algebra is that branch of mathematics which treats the common properties of algebraic systems which consist of a set, together with a reasonable notion of a 'linear combination' of elements in the set. In this section we shall define the mathematical object which experience has shown to be the most useful abstraction of this type of algebraic system.

Definition. *A* **vector space** *(or linear space) consists of the following:*

1. *a field* F *of scalars;*
2. *a set* V *of objects, called vectors;*
3. *a rule (or operation), called vector addition, which associates with each pair of vectors* α, β *in* V *a vector* $\alpha + \beta$ *in* V, *called the sum of* α *and* β, *in such a way that*

   (a) *addition is commutative,* $\alpha + \beta = \beta + \alpha$;

   (b) *addition is associative,* $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$;

(c) *there is a unique vector* 0 *in* V, *called the zero vector, such that* $\alpha + 0 = \alpha$ *for all* α *in* V;

(d) *for each vector* α *in* V *there is a unique vector* $-\alpha$ *in* V *such that* $\alpha + (-\alpha) = 0$;

4. *a rule (or operation), called scalar multiplication, which associates with each scalar* c *in* F *and vector* α *in* V *a vector* cα *in* V, *called the product of* c *and* α, *in such a way that*

(a) $1\alpha = \alpha$ *for every* α *in* V;

(b) $(c_1c_2)\alpha = c_1(c_2\alpha)$;

(c) $c(\alpha + \beta) = c\alpha + c\beta$;

(d) $(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha$.

It is important to observe, as the definition states, that a vector space is a composite object consisting of a field, a set of 'vectors,' and two operations with certain special properties. The same set of vectors may be part of a number of distinct vector spaces (see Example 5 below). When there is no chance of confusion, we may simply refer to the vector space as V, or when it is desirable to specify the field, we shall say V is a **vector space over the field** F. The name 'vector' is applied to the elements of the set V largely as a matter of convenience. The origin of the name is to be found in Example 1 below, but one should not attach too much significance to the name, since the variety of objects occurring as the vectors in V may not bear much resemblance to any preassigned concept of vector which the reader has. We shall try to indicate this variety by a list of examples; our list will be enlarged considerably as we begin to study vector spaces.

EXAMPLE 1. **The n-tuple space, F^n.** Let F be any field, and let V be the set of all n-tuples $\alpha = (x_1, x_2, \ldots, x_n)$ of scalars x_i in F. If $\beta = (y_1, y_2, \ldots, y_n)$ with y_i in F, the sum of α and β is defined by

$$\alpha + \beta = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n). \tag{2-1}$$

The product of a scalar c and vector α is defined by

$$c\alpha = (cx_1, cx_2, \ldots, cx_n). \tag{2-2}$$

The fact that this vector addition and scalar multiplication satisfy conditions (3) and (4) is easy to verify, using the similar properties of addition and multiplication of elements of F.

EXAMPLE 2. **The space of $m \times n$ matrices, $F^{m \times n}$.** Let F be any field and let m and n be positive integers. Let $F^{m \times n}$ be the set of all $m \times n$ matrices over the field F. The sum of two vectors A and B in $F^{m \times n}$ is defined by

$$(A + B)_{ij} = A_{ij} + B_{ij}. \tag{2-3}$$