5. Determinants

5.1. Commutative Rings

In this chapter we shall prove the essential facts about determinants of square matrices. We shall do this not only for matrices over a field, but also for matrices with entries which are 'scalars' of a more general type. There are two reasons for this generality. First, at certain points in the next chapter, we shall find it necessary to deal with determinants of matrices with polynomial entries. Second, in the treatment of determinants which we present, one of the axioms for a field plays no role, namely, the axiom which guarantees a multiplicative inverse for each non-zero element. For these reasons, it is appropriate to develop the theory of determinants for matrices, the entries of which are elements from a commutative ring with identity.

Definition. *A* **ring** *is a set* K, *together with two operations* $(x, y) \rightarrow x + y$ *and* $(x, y) \rightarrow xy$ *satisfying*

(a) K *is a commutative group under the operation* $(x, y) \rightarrow x + y$ (K *is a commutative group under addition*);

(b) $(xy)z = x(yz)$ (*multiplication is associative*);

(c) $x(y + z) = xy + xz$; $(y + z)x = yx + zx$ (*the two distributive laws hold*).

If $xy = yx$ *for all* x *and* y *in* K, *we say that the ring* K *is* **commutative.** *If there is an element* 1 *in* K *such that* $1x = x1 = x$ *for each* x, K *is said to be a* **ring with identity,** *and* 1 *is called the* **identity** *for* K.

We are interested here in commutative rings with identity. Such a ring can be described briefly as a set K, together with two operations which satisfy all the axioms for a field given in Chapter 1, except possibly for axiom (8) and the condition $1 \neq 0$. Thus, a field is a commutative ring with non-zero identity such that to each non-zero x there corresponds an element x^{-1} with $xx^{-1} = 1$. The set of integers, with the usual operations, is a commutative ring with identity which is not a field. Another commutative ring with identity is the set of all polynomials over a field, together with the addition and multiplication which we have defined for polynomials.

If K is a commutative ring with identity, we define an $m \times n$ matrix over K to be a function A from the set of pairs (i, j) of integers, $1 \leq i \leq m$, $1 \leq j \leq n$, into K. As usual we represent such a matrix by a rectangular array having m rows and n columns. The sum and product of matrices over K are defined as for matrices over a field

$$\begin{aligned}(A + B)_{ij} &= A_{ij} + B_{ij}\\ (AB)_{ij} &= \sum_k A_{ik}B_{kj}\end{aligned}$$

the sum being defined when A and B have the same number of rows and the same number of columns, the product being defined when the number of columns of A is equal to the number of rows of B. The basic algebraic properties of these operations are again valid. For example,

$$A(B + C) = AB + AC, \qquad (AB)C = A(BC), \qquad \text{etc.}$$

As in the case of fields, we shall refer to the elements of K as scalars. We may then define linear combinations of the rows or columns of a matrix as we did earlier. Roughly speaking, all that we previously did for matrices over a field is valid for matrices over K, excluding those results which depended upon the ability to 'divide' in K.

5.2. Determinant Functions

Let K be a commutative ring with identity. We wish to assign to each $n \times n$ (square) matrix over K a scalar (element of K) to be known as the determinant of the matrix. It is possible to define the determinant of a square matrix A by simply writing down a formula for this determinant in terms of the entries of A. One can then deduce the various properties of determinants from this formula. However, such a formula is rather complicated, and to gain some technical advantage we shall proceed as follows. We shall define a 'determinant function' on $K^{n \times n}$ as a function which assigns to each $n \times n$ matrix over K a scalar, the function having these special properties. It is linear as a function of each of the rows of the matrix; its value is 0 on any matrix having two equal rows; and its value on the $n \times n$ identity matrix is 1. We shall prove that such a function exists, and then that it is unique, i.e., that there is precisely one such function. As we prove the uniqueness, an explicit formula for the determinant will be obtained, along with many of its useful properties.

This section will be devoted to the definition of 'determinant function' and to the proof that at least one such function exists.

Definition. *Let* K *be a commutative ring with identity,* n *a positive integer, and let* D *be a function which assigns to each* n $\times$ n *matrix* A *over* K *a scalar* D(A) *in* K. *We say that* D *is* **n-linear** *if for each* i, $1 \leq i \leq n$, D *is a linear function of the* ith *row when the other* (n $-$ 1) *rows are held fixed.*

This definition requires some clarification. If D is a function from $K^{n\times n}$ into K, and if $\alpha_1, \ldots, \alpha_n$ are the rows of the matrix A, let us also write

$$D(A) = D(\alpha_1, \ldots, \alpha_n)$$

that is, let us also think of D as the function of the rows of A. The statement that D is n-linear then means

$$\text{(5-1)} \qquad \begin{aligned} D(\alpha_1, \ldots, c\alpha_i + \alpha_i', \ldots, \alpha_n) = cD(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n) \\ + D(\alpha_1, \ldots, \alpha_i', \ldots, \alpha_n). \end{aligned}$$

If we fix all rows except row i and regard D as a function of the ith row, it is often convenient to write $D(\alpha_i)$ for $D(A)$. Thus, we may abbreviate (5-1) to

$$D(c\alpha_i + \alpha_i') = cD(\alpha_i) + D(\alpha_i')$$

so long as it is clear what the meaning is.

EXAMPLE 1. Let $k_1, \ldots, k_n$ be positive integers, $1 \leq k_i \leq n$, and let a be an element of K. For each $n \times n$ matrix A over K, define

$$\text{(5-2)} \qquad D(A) = aA(1, k_1) \cdots A(n, k_n).$$

Then the function D defined by (5-2) is n-linear. For, if we regard D as a function of the ith row of A, the others being fixed, we may write

$$D(\alpha_i) = A(i, k_i)b$$

where b is some fixed element of K. Let $\alpha_i' = (A_{i1}', \ldots, A_{in}')$. Then we have

$$\begin{aligned} D(c\alpha_i + \alpha_{iz}') &= [cA(i, k_i) + A'(i, k_i)]b \\ &= cD(\alpha_i) + D(\alpha_i'). \end{aligned}$$

Thus D is a linear function of each of the rows of A.

A particular n-linear function of this type is

$$D(A) = A_{11}A_{22} \cdots A_{nn}.$$

In other words, the 'product of the diagonal entries' is an n-linear function on $K^{n\times n}$.

EXAMPLE 2. Let us find all 2-linear functions on 2×2 matrices over K. Let D be such a function. If we denote the rows of the 2×2 identity matrix by ϵ_1, ϵ_2, we have

$$D(A) = D(A_{11}\epsilon_1 + A_{12}\epsilon_2, A_{21}\epsilon_1 + A_{22}\epsilon_2).$$

Using the fact that D is 2-linear, (5-1), we have

$$\begin{aligned} D(A) &= A_{11}D(\epsilon_1, A_{21}\epsilon_1 + A_{22}\epsilon_2) + A_{12}D(\epsilon_2, A_{21}\epsilon_1 + A_{22}\epsilon_2) \\ &= A_{11}A_{21}D(\epsilon_1, \epsilon_1) + A_{11}A_{22}D(\epsilon_1, \epsilon_2) \\ &\qquad + A_{12}A_{21}D(\epsilon_2, \epsilon_1) + A_{12}A_{22}D(\epsilon_2, \epsilon_2). \end{aligned}$$

Thus D is completely determined by the four scalars

$$D(\epsilon_1, \epsilon_1), \quad D(\epsilon_1, \epsilon_2), \quad D(\epsilon_2, \epsilon_1), \quad \text{and} \quad D(\epsilon_2, \epsilon_2).$$

The reader should find it easy to verify the following. If a, b, c, d are any four scalars in K and if we define

$$D(A) = A_{11}A_{21}a + A_{11}A_{22}b + A_{12}A_{21}c + A_{12}A_{22}d$$

then D is a 2-linear function on 2×2 matrices over K and

$$\begin{aligned} D(\epsilon_1, \epsilon_1) &= a, & D(\epsilon_1, \epsilon_2) &= b \\ D(\epsilon_2, \epsilon_1) &= c, & D(\epsilon_2, \epsilon_2) &= d. \end{aligned}$$

Lemma. *A linear combination of* n*-linear functions is* n*-linear.*

Proof. It suffices to prove that a linear combination of two n-linear functions is n-linear. Let D and E be n-linear functions. If a and b belong to K, the linear combination $aD + bE$ is of course defined by

$$(aD + bE)(A) = aD(A) + bE(A).$$

Hence, if we fix all rows except row i

$$\begin{aligned} (aD + bE)(c\alpha_i + \alpha_i') &= aD(c\alpha_i + \alpha_i') + bE(c\alpha_i + \alpha_i') \\ &= acD(\alpha_i) + aD(\alpha_i') + bcE(\alpha_i) + bE(\alpha_i') \\ &= c(aD + bE)(\alpha_i) + (aD + bE)(\alpha_i'). \quad \blacksquare \end{aligned}$$

If K is a field and V is the set of $n \times n$ matrices over K, the above lemma says the following. The set of n-linear functions on V is a subspace of the space of all functions from V into K.

EXAMPLE 3. Let D be the function defined on 2×2 matrices over K by

$$D(A) = A_{11}A_{22} - A_{12}A_{21}. \tag{5-3}$$

Now D is the sum of two functions of the type described in Example 1:

$$\begin{aligned} D &= D_1 + D_2 \\ D_1(A) &= A_{11}A_{22} \\ D_2(A) &= -A_{12}A_{21}. \end{aligned}$$

By the above lemma, D is a 2-linear function. The reader who has had any experience with determinants will not find this surprising, since he will recognize (5-3) as the usual definition of the determinant of a 2×2 matrix. Of course the function D we have just defined is not a typical 2-linear function. It has many special properties. Let us note some of these properties. First, if I is the 2×2 identity matrix, then $D(I) = 1$, i.e., $D(\epsilon_1, \epsilon_2) = 1$. Second, if the two rows of A are equal, then

$$D(A) = A_{11}A_{12} - A_{12}A_{11} = 0.$$

Third, if A' is the matrix obtained from a 2×2 matrix A by interchanging its rows, then $D(A') = -D(A)$; for

$$\begin{aligned} D(A') &= A'_{11}A'_{22} - A'_{12}A'_{21} \\ &= A_{21}A_{12} - A_{22}A_{11} \\ &= -D(A). \end{aligned}$$

Definition. *Let* D *be an* n-*linear function. We say* D *is* **alternating** (*or* **alternate**) *if the following two conditions are satisfied:*

(a) D(A) = 0 *whenever two rows of* A *are equal.*

(b) *If* A′ *is a matrix obtained from* A *by interchanging two rows of* A, *then* D(A′) = −D(A).

We shall prove below that any n-linear function D which satisfies (a) automatically satisfies (b). We have put both properties in the definition of alternating n-linear function as a matter of convenience. The reader will probably also note that if D satisfies (b) and A is a matrix with two equal rows, then $D(A) = -D(A)$. It is tempting to conclude that D satisfies condition (a) as well. This is true, for example, if K is a field in which $1 + 1 \neq 0$, but in general (a) is not a consequence of (b).

Definition. *Let* K *be a commutative ring with identity, and let* n *be a positive integer. Suppose* D *is a function from* n × n *matrices over* K *into* K. *We say that* D *is a* **determinant function** *if* D *is* n-*linear, alternating, and* D(I) = 1.

As we stated earlier, we shall ultimately show that there is exactly one determinant function on $n \times n$ matrices over K. This is easily seen for 1×1 matrices $A = [a]$ over K. The function D given by $D(A) = a$ is a determinant function, and clearly this is the only determinant function on 1×1 matrices. We are also in a position to dispose of the case $n = 2$. The function

$$D(A) = A_{11}A_{22} - A_{12}A_{21}$$

was shown in Example 3 to be a determinant function. Furthermore, the formula exhibited in Example 2 shows that D is the only determinant