NON-SQUARE DETERMINANTS AND MULTILINEAR VECTORS

H. RANDOLPH PYLE, Whittier College

When a student learns of the many relationships between determinants and square matrices, he is apt to wonder why there isn't a theory of non-square determinants to go with that of non-square matrices. As a matter of fact it is quite possible to present such a theory as we shall see. In order to keep the length of the presentation to a minimum we shall base our discussion on that of a standard modern text, namely Chapter 5 of "Elements of Linear Algebra" by L. J. Paige and J. D. Swift [1].

Vectors in the real vector space $V_n(R)$ are defined as sets of ordered n-tuples of real numbers, $X = (x_1, x_2, \cdots, x_n)$. A determinant $D(X_1, X_2, \cdots, X_n)$ is defined as a real number assigned to the ordered set of n vectors $X_1, X_2, \cdots, X_n$, which has the following properties: (1) D is linear in each of its arguments, (2) D is zero if any two of its arguments are equal, (3) $D = 1$ when $X_i = E_i$, where E_i is the natural basis vector whose ith component is one and whose other components are zero.

The classical properties of determinants are developed from this definition. The fact that we have n vectors, each with n components, gives us the square array of the determinant.

Now suppose that we revise our definition so that we have r vectors ($r \leqq n$) instead of n vectors, and reword our definition as follows:

Definition: A determinant $D(X_1, \cdots, X_r)$, $r \leqq n$, is an ordered set of $\binom{n}{r}$ real numbers corresponding to the ordered set of r vectors $X_1, \cdots, X_r$ in $V_n(R)$, which has the properties listed above.

This ordered set of $\binom{n}{r}$ real numbers consists of the numbers associated with each of the $\binom{n}{r}$ square r-rowed determinants that can be formed by selecting r columns from the array of vectors, arranged in a predetermined order. For example, if $n = 3$, $r = 2$, and $X_i = (x_{i1}, x_{i2}, x_{i3})$, the (2×3)-determinant

$$D(X_1, X_2) = \begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{vmatrix} = \left(\begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix}, \begin{vmatrix} x_{11} & x_{13} \\ x_{21} & x_{23} \end{vmatrix}, \begin{vmatrix} x_{12} & x_{13} \\ x_{22} & x_{23} \end{vmatrix} \right)$$

Since each of the components of the $(r\times n)$-determinant involves the r rows of the determinant, any valid operation on the rows of a square determinant is equally valid for each component of the non-square determinant. For instance, if two rows are permuted the sign of each component of D is changed and we can say that D changes sign. If the row vectors are linearly dependent, each component of D vanishes and we say that $D = 0$. Conversely, if $D = 0$, the vectors are dependent. If every element of a row is zero, or if two rows are proportional, $D = 0$. If a multiple of one row is added to another row, the value of the determinant is not changed.

But it is no longer true that the transpose of a determinant equals the determinant. The usual theorem that the roles of columns and rows may be interchanged no longer holds. Instead we may say that any operation valid for the rows of a determinant is valid for the columns of its transpose.

In defining equality and the sum, product and magnitude of non-square determinants we may treat them as vectors in an Euclidean space of $\binom{n}{r}$ dimensions. Two non-square determinants are equal when their corresponding components are equal.

If

$$D_1(X_1, \cdots, X_r) = (u_1, u_2, \cdots, u_k) ,$$

$k = \binom{n}{r}$, and

$$D_2(Y_1, \cdots, Y_r) = (v_1, v_2, \cdots, v_k) ,$$

we define the sum of D_1 and D_2 as a new vector whose components are the sums of corresponding components of D_1 and D_2, i. e.,

$$D_1 + D_2 = (u_1 + v_1, u_2 + v_2, \cdots, u_k + v_k) .$$

The product of D_1 and D_2 is defined as the vector inner product, i. e., the sum of the products of corresponding components. We shall designate it by $D_1 \cdot D_2$, so that

$$D_1 \cdot D_2 = u_1 v_1 + u_2 v_2 + \cdots + u_k v_k .$$

The magnitude of D_1 is defined as the square root of the sum of the squares of its components, so that

$$|D_1| = (D_1 \cdot D_1)^{1/2} = (u_1^2 + u_2^2 + \cdots + u_k^2)^{1/2} .$$

An alternative approach to the product is to use matrix multiplication. If we designate the corresponding matrices as

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_r \end{pmatrix}, \quad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_r \end{pmatrix}$$

where the X_i, Y_i are the row vectors of the matrices, the transpose of Y is $Y^t = (Y_1^t, Y_2^t, \cdots, Y_r^t)$, where the Y_i^t are the column vectors which are transposes of the row vectors of Y. The matrix product XY^t is square and its determinant is an orthodox one. It can be shown [2] that this determinant has the same value as $D_1 \cdot D_2$ already defined. That is

$$D_1 \cdot D_2 = u_1 v_1 + u_2 v_2 + \cdots + u_k v_k = \begin{vmatrix} X_1 \cdot Y_1 & X_1 \cdot Y_2 & \cdots & X_1 \cdot Y_r \\ X_2 \cdot Y_1 & X_2 \cdot Y_2 & \cdots & X_2 \cdot Y_r \\ & \cdot \cdot \cdot & & \\ X_r \cdot Y_1 & X_r \cdot Y_2 & \cdots & X_r \cdot Y_r \end{vmatrix}$$

The latter form of the product is frequently best for computation.

When $X_i = Y_i$, $D_1 = D_2$ and the magnitude of D is

$$|D| = (D \cdot D)^{1/2} = \begin{vmatrix} X_1 \cdot X_1 & X_1 \cdot X_2 & \cdots & X_1 \cdot X_r \\ & \cdots & & \\ X_r \cdot X_1 & X_r \cdot X_2 & \cdots & X_r \cdot X_r \end{vmatrix}^{1/2}.$$

Content of a parallelotope. Paige and Swift define the content (volume) of a parallelotope with the n edges $X_1, \cdots, X_n$ in Euclidean space of n dimensions as the magnitude of the determinant $D(X_1, \cdots, X_n)$. We can generalize this to say that the content of a parallelotope with r edges $X_1, \cdots, X_r$ in Euclidean n-space ($r \leqq n$) is equal to the magnitude of the non-square determinant $D(X_1, \cdots, X_r)$. This reduces to the usual formulas in two and three dimensions. For instance, the area of a parallelogram in 3-space whose edges are the vectors X_1, X_2 may be found by taking the magnitude of the cross product of X_1 and X_2. When $X_i = (x_{i1}, x_{i2}, x_{i3})$,

$$X_1 \times X_2 = \left(\begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix}, \begin{vmatrix} x_{13} & x_{11} \\ x_{23} & x_{21} \end{vmatrix}, \begin{vmatrix} x_{12} & x_{13} \\ x_{22} & x_{23} \end{vmatrix} \right).$$

Except for the sign of the middle term, these components are the same as those of $D(X_1, X_2)$ so that $|D(X_1, X_2)|$ gives the area of the parallelogram.

Multilinear vectors. Grassmann, Cartan, Bourbaki and other writers have named these parallelotopes multilinear vectors. Cartan [3] calls a parallelogram a bivector and a parallelepiped a trivector. They represent the outer product of the line vectors that form their edges. He uses brackets $[X_1X_2\cdots X_r]$ for this product and calls it an r-vector. It is best represented by an $(r\times n)$-determinant. The components of the determinant are the coordinates of the r-vector in n-space. The magnitude of the r-vector is the magnitude of the determinant. Two r-vectors are equal when their determinants are equal. The sum of two such vectors is the sum of their determinants and their scalar product is the product of the determinants. This scalar product is the magnitude of the first times the projection of the second on the space spanned by the first. If we designate the r-vectors by $A_{(r)}$ and $B_{(r)}$, where the subscript in parentheses represents the dimension of the vector, and write their scalar product as $A_{(r)} \cdot B_{(r)}$, we define $A_{(r)} \cdot B_{(r)} = |A_{(r)}||B_{(r)}| \cos \theta$, where θ is the angle between the two r-flats. This serves to define θ.

Two r-vectors are said to be orthogonal when their inner product is zero. This means that at least one line vector in each is perpendicular to every line vector in the other. To show this let $L = x_1B_1 + x_2B_2 + \cdots + x_rB_r$, where the B_i are line vectors defining $B_{(r)}$ and the x_i are scalars. If L is orthogonal to every vector in $A_{(r)}$, $A_i \cdot L = 0$, $(i = 1, \cdots, r)$. This gives a set of r homogeneous equations in the r variables $x_1, \cdots, x_r$, whose coefficients are of the form $A_i \cdot B_j$. These equations will have a non-trivial solution if and only if the determinant of the coefficients vanishes. But this determinant is $A_{(r)} \cdot B_{(r)}$, so that the vanishing of the inner product assures such a vector.

When every line vector in $B_{(r)}$ is perpendicular to every line vector in $A_{(r)}$, we say that the two r-vectors are completely orthogonal. Then every element in the determinant of the product vanishes.

Reciprocal multiple vectors and determinants. Greville [4] discusses the pseudoinverse, as he calls it, of a non-square matrix. If we have an r-vector $A_{(r)} = [A_1 A_2 \cdots A_r]$, the set of line vectors A_i^* reciprocal to the vectors A_i with respect to the r-flat define an r-vector $A_{(r)}^* = [A_1^* A_2^* \cdots A_r^*]$, whose matrix turns out to be the pseudoinverse of Greville. In vector terms $A_{(r)} \cdot A_{(r)}^* = 1$, and $A_{(r)}$ and $A_{(r)}^*$ are reciprocal vectors. We may call their determinants reciprocal $(r \times n)$-determinants.

$A_{(r)}^*$ is easily found by matrix methods. If $(A_{(r)}^2)^{-1}$ is the inverse of the square matrix corresponding to $A_{(r)} \cdot A_{(r)}$, then using matrix multiplication, $A_{(r)}^* = (A_{(r)}^2)^{-1} A_{(r)}$.

There are many interesting relations in the theory of multilinear vectors, but these are enough to show how they relate to non-square determinants.

Relations between the components of an $(r \times n)$-determinant. We have said that the $\binom{n}{r}$ components of an $(r \times n)$-determinant resemble the components of a vector in a space of $\binom{n}{r}$ dimensions. This is true but there are identities between the components that limit the number of independent ones. In order to find these identities, let us consider

$$D(X_1, \cdots, X_r) = \begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ & & \cdots & \\ x_{r1} & x_{r2} & \cdots & x_{rn} \end{vmatrix}, \quad r \leq n .$$

If we form an $(r+1)$-rowed determinant by putting any one of the vectors $X_1, \cdots, X_r$ at the top of the existing determinant, the new matrix is still of rank r (assuming that the original vectors are independent) because the new vector is a repetition of one of the old ones. This means that every $(r+1)$-rowed square determinant in the array vanishes. Expanding by the elements of the top row gives an identity relation between the elements of the vector in that row and some of the components of $D(X_1, \cdots, X_r)$.

There will be $\binom{n}{r+1}$ identities for each vector added at the top and r

such vectors that may be used, making a total of $r\binom{n}{r+1}$ identities. This procedure can be repeated with two rows added at the top, giving $\binom{r}{2}\binom{n}{r+1}$ identities, involving the components of $D(X_1, \cdots, X_r)$ and the 2-rowed minors of the elements in the top two rows. This process can be repeated until we add r rows.

When $n \geqq 2r$ we can add all r vectors of the original determinant. Then the identities can be expressed in terms of the components of D only. There will be $\binom{n}{2r}$ of them. For example, when $n = 4$, $r = 2$,

$$D(X_1, X_2) = \begin{vmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \end{vmatrix} = (u_{12}, u_{13}, u_{14}, u_{23}, u_{24}, u_{34})$$

where the subscripts of the u's indicate the columns involved in the component determinants. The augmented determinants are

$$\begin{vmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \end{vmatrix}, \quad \begin{vmatrix} x_{21} & x_{22} & x_{23} & x_{24} \\ x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \end{vmatrix}, \quad \begin{vmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \end{vmatrix}$$

From the first of these we get the identities:

$$x_{11}u_{23} - x_{12}u_{13} + x_{13}u_{12} = 0 \qquad x_{11}u_{34} - x_{13}u_{14} + x_{14}u_{13} = 0$$

$$x_{11}u_{24} - x_{12}u_{14} + x_{14}u_{12} = 0 \qquad x_{12}u_{34} - x_{13}u_{24} + x_{14}u_{23} = 0 .$$

From the second we get four more derived from these by replacing x_{1i} by x_{2i}. From the last determinant we have

$$u_{12}u_{34} - u_{13}u_{24} + u_{14}u_{23} = 0 .$$

Conclusion. When presented in this fashion the theory of non-square determinants seems to follow naturally from that of square determinants. Their applications are useful extensions of those of square determinants.

REFERENCES

1. Paige and Swift, *Elements of Linear Algebra*, Ginn and Co., 1961.
2. Thrall and Tornheim, *Vector Spaces and Matrices*, John Wiley and Sons, 1957, p. 127.
3. E. Cartan, *Lecons sur la Geometrie des Espaces de Riemann,* Paris, 1928.
4. T. N. E. Greville, *The Pseudoinverse of a Rectangular or Singular Matrix and its Application to the Solution of Systems of Linear Equations*, SIAM Newsletter, Vol. V, Number 2, March 1957.
5. Muir and Metzler, *Theory of Determinants*, Albany, N. Y., 1930.
6. Bourbaki, *Elements de Mathematique, Algebre Multilineare*, Paris, 1958.