ON IDENTITIES ASSOCIATED WITH A DISCRIMINANT

by M. J. NEWELL

(Received 26th April 1972, revised 22nd October 1972)

1. Introduction

If the elements of a symmetric matrix lie in the real field it is well known that the roots of its characteristic equation are real. This implies that the discriminant of that equation (i.e. the product of the squared differences of the roots) is a polynomial in the elements which is non-negative and the same must be true for the leading coefficients of all the other Sturm functions associated with the characteristic equation. One would expect that it should be possible to express them as a sum of squares. Conversely, such an expression would establish the reality of the roots.

The discriminant of a symmetric matrix of order three has been considered by (**1**) Kummer (1843), (**2**) Tannery (1883), (**3**) Watson (1956). The last two authors established by different methods such an identity first given by Kummer who, however, gave no indication of the method used. In this paper it is shown that a similar identity exists when the order is n and Kummer's identity is then derived for the case $n = 3$. The extension of these results to Hermitian matrices is immediate if, as usual, square is replaced by squared modulus.

2. Preliminary result

If the elements of any real square matrix X of order n are written without repetition in one row and in the order in which they appear in the successive rows of X we may denote this arrangement of n^2 elements by

$$x_{11}, x_{12}, \ldots, x_{rs}, \ldots, x_{nn}.$$

Similarly we denote by $x_{rs}^{(k)}$ the typical element of the corresponding arrangement for the matrix X^k which is the kth power of X.

With the n sequences of n^2 terms each which arise in this way from the matrices

$$X^0 = I, X, X^2, \ldots, X^{n-1},$$

we form a matrix M having n rows and n^2 columns.

If we had arranged the elements of X in the order in which they appear in the successive columns of X and if we had written the n sequences arising from the matrices $I, X, X^2, \ldots, X^{n-1}$ as columns rather than rows we could construct another matrix N having n^2 rows and n columns.

https://doi.org/10.1017/S0013091500010063 Published online by Cambridge University Press

Note that $x_{rs}^{(k)}$ lies in the $(k+1)$th row and in the $[n(r-1)+s]$th column of M. On the other hand it is the element $x_{sr}^{(k)}$ which lies in the $(k+1)$th column and $[n(r-1)+s]$th row of N. Accordingly if in the matrix M transposed we replace each element $x_{rs}^{(k)}$ by $x_{sr}^{(k)}$ we arrive at the matrix N.

From the fact that the trace of the product XY of two matrices (i.e. the sum of its diagonal elements) is $\sum_{rs} x_{rs}y_{sr}$ it will be seen that the product MN is a square matrix of order n in which the element lying in the rth row and sth column is tr (X^{r+s-2}).

If the latent roots of X are $\lambda_1, \lambda_2, \ldots, \lambda_n$, it is well known that

$$\operatorname{tr}(X^k) = \Sigma\lambda_p^k.$$

Consequently the product MN is a matrix Z where

$$Z_{rs} = \sum_{p=1}^{n} \lambda_p^{r+s-2} \quad (r, s = 1, 2, \ldots, n).$$

If, however, the simple alternant matrix A of the variables λ_p is written down, i.e. $a_{rs} = \lambda_s^{r-1}$, it will be seen that A multiplied by its transpose A' is also equal to Z, i.e.

$$MN = AA'. \tag{1}$$

The matrix AA' is the matrix in which the element lying in the rth row and sth column is S_{r+s-2}, where $S_k = \Sigma\lambda_p^k$ and the determinant of the matrix AA' (being equal to the square of det A) is $\prod_{r<s}(\lambda_r-\lambda_s)^2$ which is (by definition) the discriminant of the characteristic equation of X. Thus

$$\prod_{r<s}(\lambda_r-\lambda_s)^2 = \det(MN). \tag{2}$$

3. Symmetric matrices

If the general matrix X in the above discussion is replaced by the symmetric matrix S (i.e. $x_{rs} = x_{sr}$), then every power of S is also symmetric and it is evident that the matrix N is the transpose M' of the matrix M. When det (MM') in (2) is expanded in accordance with a well-known theorem on the multiplication of determinantal arrays (see **4**, p. 93) we have an expression for the discriminant as the sum of the squares of all the n-rowed minors of M.

The same theorem shows that the leading principal r-rowed minor of AA' can be written as a sum of squares of the appropriate r-rowed minors of M, and each of these leading principal minors is well known (**5**, p. 333) to be the leading coefficient of the corresponding member of Sturm's auxiliary functions, multiplied by a squared factor.

If X is replaced by an Hermitian matrix H, i.e. x_{rs} and x_{sr} are complex conjugates, then the matrix N is the Hermitian conjugate M^* and we get the sum of the squared moduli of all the n-rowed minors of M.

This is sufficient to establish the reality of the latent roots in both cases.

https://doi.org/10.1017/S0013091500010063 Published online by Cambridge University Press

4. Reduction of the number of items

Returning to the case of the symmetric matrix S we see that every column of M corresponding to a distinct pair of suffixes r, s is repeated under the column corresponding to s, r. Since we are concerned with non-zero minors only we may restrict attention to the $\frac{1}{2}n(n+1)$ columns designated by r, s where $r \leqq s$ provided that the square of every n-rowed minor is supplied with a factor 2 for each such column with $r<s$ appearing in it in order to allow for the fact that the minor occurs again when we choose the column sr instead of rs.

Also, from the structure of the first row of M it can be seen that every non-zero minor must have at least one column corresponding to a " diagonal " pair rr. It is convenient to adjust the matrix M by putting the columns corresponding to 11, 22, ..., nn, in the first n places and by suppressing the columns rs where $r>s$. We say that an n-rowed minor is of class x if there are x terms equal to unity (and consequently $n-x$ zeros) in its first row.

We make use of certain linear relations among the minors which stem from the following two considerations.

Firstly, taking a minor of class x, if it does not contain the column designated by a specified double suffix (e.g. the first column 11) we can form a determinant of order $n+1$ by adjoining this first column and repeating the first row. The expansion of this vanishing determinant in terms of its first row will express the minor concerned as a linear function (with coefficients ± 1) of other minors of this class, each of which contains the first column and in each of which the $n-x$ columns corresponding to the non-equal suffixes are undisturbed. Hence for a fixed choice of the last $n-x$ suffixes the ${}^nC_x$ n-rowed minors containing them are not independent and each of them can be written as a linear combination of the subset of ${}^{n-1}C_{x-1}$ such minors which contain the first column (or indeed any other specified column s, s).

Thus, for example, there is one linear relation between the n minors of class $n-1$ which incorporate one fixed column of type rs.

Apart from these relations between minors of the same class there is another set, connecting minors of different classes, which may be obtained as follows.

For any integral values of j and k the product S^jS^k of powers of the symmetric matrix S is itself symmetric so that the element in the pth row and qth column of the product viz. $\sum_t S_{pt}^{(j)}S_{tq}^{(k)}$ is equal to the element in the qth row and pth column, viz. $\sum_u S_{qu}^{(j)}S_{up}^{(k)}$. Written out in full this equation is

$$S_{p1}^{(j)}S_{1q}^{(k)}+S_{p2}^{(j)}S_{2q}^{(k)}+\ldots S_{pp}^{(j)}S_{pq}^{(k)}+\ldots S_{pq}^{(j)}S_{qq}^{(k)}\ldots S_{pn}^{(j)}S_{nq}^{(k)}$$
$$= S_{q1}^{(j)}S_{1p}^{(k)}+S_{q2}^{(j)}S_{2p}^{(k)}+\ldots S_{qp}^{(j)}S_{pp}^{(k)}+\ldots S_{qq}^{(j)}S_{qp}^{(k)}\ldots S_{qn}^{(j)}S_{np}^{(k)}.$$

Paying special attention to the four terms in which the repeated suffixes pp or qq occur we may write the equation in the form

$$(S_{qq}^{(j)}-S_{pp}^{(j)})S_{pq}^{(k)}-S_{pq}^{(j)}(S_{qq}^{(k)}-S_{pp}^{(k)}) = \sum_t{}' \begin{vmatrix} S_{pt}^{(j)} & S_{qt}^{(j)} \\ S_{pt}^{(k)} & S_{qt}^{(k)} \end{vmatrix}, \tag{3}$$

https://doi.org/10.1017/S0013091500010063 Published online by Cambridge University Press

where Σ' means that the values $t = p$ and q are excluded in the summation of determinants on the right. The left side of the equation can also be put in the determinantal form

$$\begin{vmatrix} 1 & 1 & 0 \\ S_{pp}^{(j)} & S_{qq}^{(j)} & S_{pq}^{(j)} \\ S_{pp}^{(k)} & S_{qq}^{(k)} & S_{pq}^{(k)} \end{vmatrix}.$$

This relation holds for all integral values of $p, q<r$ and may be used as follows. Whenever the special combination of three columns pp, qq and pq occurs in any n-rowed minor of class 2 a consideration of its Laplace expansion in terms of the minors from these columns will show that the three columns may be replaced by the columns pt and qt provided that we sum for all values of t excepting $t = p$ or q. The process reduces the class of the n-rowed minor by 2.

We can now deal rather easily with the case n = 3.

5. Kummer's Identity

If $S^3 - p_1S^2 + p_2S - p_3I = 0$ is the characteristic equation of S, we have

$$S(S^2 - p_1S + p_2I) = (\det S)I,$$

so that the matrix adj S formed from the co-factors of S is the second factor on the left side of this equation. In the matrix corresponding to M•(Section 2) it is convenient to replace S^2 by adj S. This has the effect of adding to the last row a certain linear combination of the preceding rows and leaves all three-rowed minors unaltered.

$$\begin{matrix} 1 & 1 & 1 & \cdot & \cdot & \cdot \\ a_{11} & a_{22} & a_{22} & a_{12} & a_{13} & a_{23} \\ A_{11} & A_{22} & A_{33} & A_{12} & A_{13} & A_{23}. \end{matrix}$$

There is only one minor of class three which, using an easily understandable notation, may be written

$$| a_{11} \quad a_{22} \quad a_{33} |.$$

Three minors of class two contain the column 12 and from the relations in Section 4 we see that one of them reduces, viz.

$$| a_{11} \quad a_{22} \quad a_{12} | = | a_{13} \quad a_{23} |,$$

and that the linear relation connecting all three can be written

$$| a_{22} \quad a_{33} \quad a_{12} | - | a_{11} \quad a_{33} \quad a_{12} | = -| a_{13} \quad a_{23} |.$$

The squares of each of these minors must be given a coefficient 2 and using the identity $2x^2+2y^2 = (x+y)^2+(x-y)^2$, we may reduce the three terms to two, viz.

$$\{| a_{11} \quad a_{33} \quad a_{12} | + | a_{22} \quad a_{33} \quad a_{12} |\}^2 + 3| a_{13} \quad a_{23} |^2.$$

https://doi.org/10.1017/S0013091500010063 Published online by Cambridge University Press

There are similar expressions arising from the minors of class 2 which contain the columns 13 or 23.

Finally, there are three equal minors of class one which contain the columns 12 and 13 each having a coefficient 2^2 and giving a term $12 \mid a_{12}\, a_{13} \mid^2$ and there are similar results arising from the other two pairs of columns 12, 23 and 13, 23. Collecting terms we find that the right-hand side of the equation (2) can be put in the form

$$\mid a_{11} \quad a_{22} \quad a_{33} \mid^2 + \Sigma\{\mid a_{11} \quad a_{33} \quad a_{12} \mid + \mid a_{22} \quad a_{33} \quad a_{12} \mid\}^2 + 15\Sigma \mid a_{12} \quad a_{13} \mid^2. \quad (4)$$

The left side (2) is a well-known polynomial in p_1, p_2, p_3 and the equation of the two forms is Kummer's identity.

REFERENCES

(1) E. E. Kummer, *Journal für Math.* **26** (1843), 268-272.

(2) J. Tannery, *Bulletin des Sci. Math. et Astr.* (2), **7** (1883), 103-107.

(3) G. N. Watson, On identities associated with a discriminant. *Proc. Edinburgh Math. Soc.* **10** (1956), 101-107.

(4) A. C. Aitken, *Determinants and Matrices* (Oliver and Boyd, 1939).

(5) Burnside and Panton, *Theory of Equations* (Dublin Univ. Press (Third Ed.), 1892).

University College
Galway, Ireland