THE INVARIANCE OF SYMMETRIC FUNCTIONS OF SINGULAR VALUES

MARVIN MARCUS AND HENRYK MINC

Let $M_{m,n}$ denote the vector space of all $m \times n$ matrices over the complex numbers. A general problem that has been considered in many forms is the following: suppose $\mathfrak{A}$ is a subset (usually subspace) of $M_{m,n}$ and let f be a scalar valued function defined on $\mathfrak{A}$. *Determine the structure of the set* $\mathfrak{A}_f$ *of all linear transformations* T *that satisfy*

$$f(T(A)) = f(A) \text{ for all } A \in \mathfrak{A} . \tag{1}$$

The most interesting choices for f are the classical invariants such as rank [3, 4, 7] determinant [1, 2, 3, 5, 10] and more general symmetric functions of the characteristic roots [6, 8]. In case $\mathfrak{A}$ is the set of n-square real skew-symmetric matrices ($m = n$) and $f(A)$ is the Hilbert norm of A then Morita [9] proved the following interesting result: $\mathfrak{A}_f$ *consists of transformations* T *of the form*

$$T(A) = U'AU \text{ for } n \neq 4 ,$$
$$T(A) = U'AU \text{ or } T(A) = U'A^+U \text{ for } n = 4$$

where U *is a fixed real orthogonal matrix and* A^+ *is the matrix obtained from* A *by interchanging its* $(1, 4)$ *and* $(2, 3)$ *elements.*

Recall that the Hilbert norm of A is just the largest *singular value* of A (i.e., the largest characteristic root of the nonnegative Hermitian square root of A^*A).

In the present paper we determine $\mathfrak{A}_f$ when $\mathfrak{A}$ is all of $M_{m,n}$ and f is some particular elementary symmetric function of the squares of the singular values. We first introduce a bit of notation to make this statement precise. If $A \in M_{n,n}$ then $\lambda(A) = (\lambda_1(A), \cdots, \lambda_n(A))$ will denote the n-tuple of characteristic roots of A in some order. The rth elementary symmetric function of the numbers $\lambda(A)$ will be denoted by $E_r[\lambda(A)]$; this is, of course, the same as the sum of all r-square principal subdeterminants of A. We also denote by $\rho(A)$ the rank of A.

THEOREM. *A linear transformation* T *of the space* $M_{m,n}$ *leaves invariant the* r*th elementary symmetric function of the squares of the singular values of each* $A \in M_{m,n}$, *for some fixed* r, $1 < r \leqq n$, *if and only if there exist unitary matrices* U *and* V *in* $M_{m,m}$ *and* $M_{n,n}$ *respectively such that*

Received January 16, 1961. The work of the first author was supported in part by the the Office of Naval Research.

(2) $$T(A) = UAV \text{ if } m \neq n \text{ and}$$

(3) $$T(A) = UAV \text{ or } T(A) = UA'V \text{ if } m = n\ .$$

The sufficiency of (2) and (3) is clear and we prove the necessity in a sequence of lemmas some of which may be of interest in themselves. Assume without loss of generality that $m \geqq n$.

LEMMA 1. *Let $A, B \in M_{m,n}$ and let $\varphi_B(x) = E_r[\lambda((xA + B)^* (xA + B))]$ where x is a real indeterminate. Then*

(4) $$\deg \varphi_B(x) \leqq 2 \text{ for all } B \in M_{m,n}$$

if and only if

(5) $$\rho(A) \leqq 1\ .$$

Proof. We first remark that $\varphi_B(x)$ is actually a polynomial in x since it is the sum of all $\binom{n}{r}$ r-square principal subdeterminants of $(xA + B)^* (xA + B)$. The matrix A can be written, by a slight extension of the polar factorization theorem to rectangular matrices, in the form $A = UH$ where H is n-square hermitian positive semi-definite and $U \in M_{m,n}$ satisfies $U^*U = I_n$, the n-square identity matrix. Then

$$\begin{aligned}\varphi_B(x) &= E_r[\lambda((xUH + B)^* (xUH + B))] \\ &= E_r[\lambda((xH + U^*B)^* (xH + U^*B))]\ .\end{aligned}$$

Now let $H = V^*DV$ where V is unitary and D is diagonal. Then

$$\begin{aligned}\varphi_B(x) &= E_r[\lambda(V^*(xD + VU^*BV^*)^* VV^*(xD + VU^*BV^*)V)] \\ &= E_r[\lambda((xD + B_1)^* (xD + B_1))]\end{aligned}$$

where $B_1 = VU^*BV^*$. Now suppose $\rho(A) = \rho(D) = 1$. Then D has exactly one nonzero entry which we may clearly assume to be in the (1, 1) position. It follows that $(xD + B_1)^* (xD + B_1)$ has a quadratic polynomial in x in the (1, 1) position, first degree polynomials in the other first row and first column positions and constants elsewhere. Therefore, every principal subdeterminant of this matrix is a polynomial in x of degree at most 2.

On the other hand, if (4) holds then in particular for $B = 0$

$$\varphi_0(x) = E_r[\lambda(x^2D^*D)]$$

and $\deg \varphi_0(x) \leqq 2$; this implies that the diagonal matrix D^*D can have at most one nonzero entry. But then $1 \geqq \rho(D^*D) = \rho(D) = \rho(\mathrm{A})$.

LEMMA 2. *Let $f(t_1, \cdots, t_n)$ be a monotone strictly increasing function of each t_j for $t_j > 0$. If T is a linear map of $M_{m,n}$ into itself satisfying*

$$f(\lambda(A^*A)) = f(\lambda((T(A))^*T(A))), \quad A \in M_{m,n}$$

then T is nonsingular.

Proof. Suppose $T(A) = 0$. Then

$$\begin{aligned} f(\lambda(X^*X)) &= f(\lambda((T(X))^*T(X))) \\ &= f(\lambda((T(A+X))^*T(A+X))) \\ &= f(\lambda((A+X)^*(A+X))) . \end{aligned}$$

Let $A = UH$ where $U^*U = I_n$ and H is nonnegative Hermitian. Then taking $H = V^*DV$ where D is diagonal and V is unitary we find as in Lemma 1 that

$$f(\lambda(X^*X)) = f(\lambda((D+Y)^*(D+Y))) ,$$

$Y = VU^*XV^*$. Now as X runs over $M_{m,n}$ Y runs over $M_{n,n}$ and moreover

$$\lambda(X^*X) = \lambda(V^*Y^*VU^*UV^*YV) = \lambda(Y^*Y) .$$

Hence

(6) $$f(\lambda(Y^*Y)) = f(\lambda((D+Y)^*(D+Y)))$$

for all $Y \in M_{n,n}$. Let Y be a real diagonal matrix with diagonal elements $y_1, \cdots, y_n$. Then if D has diagonal elements $d_1, \cdots, d_n$ we conclude from (6) that

$$f(y_1^2, \cdots, y_n^2) = f(d_1^2 + y_1^2, \cdots, d_n^2 + y_n^2) .$$

Thus $D = 0, A = 0$ and T is nonsingular.

We remark at this point that the elementary symmetric functions satisfy the conditions of Lemma 2 and hence the T of the theorem is nonsingular.

LEMMA 3. *If $\rho(A) = 1$ then $\rho(T(A)) = 1$.*

Proof. If $\rho(A) = 1$ then, by Lemma 1, deg $\varphi_B(x) \leqq 2$. Now

$$\begin{aligned} \varphi_B(x) &= E_r[\lambda((xA+B)^*(xA+B))] \\ &= E_r[\lambda((T(xA+B))^*T(xA+B))] \\ &= E_r[\lambda((xT(A)+T(B))^*(xT(A)+T(B)))] . \end{aligned}$$

By Lemma 2 T is nonsingular so $T(B)$ ranges over $M_{m,n}$ as B does. Hence, by Lemma 1, $\rho(T(A)) \leqq 1$. But $T(A) \neq 0$ since $\rho(A) = 1$. Thus $\rho(T(A)) = 1$.

At this point we invoke [7: p. 1219] that tells us that a linear transformation on $M_{m,n}$ which preserves rank 1 has the desired form:

$$T(A) = UAV \text{ for all } A \in M_{m,n}$$

or

$$T(A) = UA'V \text{ for all } A \in M_{m,n} ,$$

where U and V are nonsingular m-square and n-square matrices respectively and the second eventuality occurs only if $m = n$. The proof of the theorem will be complete if we show

LEMMA 4. *U and V may be chosen to be unitary.*

Proof. We show this when T has the form (2); if T has the form (3) the argument is essentially the same. Let $V = HP$ and $U = QK$ where H and K are positive definite Hermitian and P and Q are unitary. Then

$$\begin{aligned} E_r[\lambda(A^*A)] &= E_r[\lambda((UAV)^*(UAV))] \\ &= E_r[\lambda(V^*A^*U^*UAV)] \\ &= E_r[\lambda(P^*HA^*K^2AHP)] \\ &= E_r[\lambda(HA^*K^2AH)] \\ &= E_r[\lambda(H^2A^*K^2A)] \end{aligned}$$

for all A. Let $H = XDX^*, K = YGY^*$, X and Y unitary, D and G diagonal matrices with main diagonals $d_1, \cdots, d_n$ and $g_1, \cdots, g_n$ respectively. Then

$$\begin{aligned} E_r[\lambda(A^*A)] &= E_r[\lambda(XD^2X^*A^*YG^2Y^*A)] \\ &= E_r[\lambda(D^2B^*G^2B)] \end{aligned}$$

where $B = Y^*AX$. Now

$$\lambda(A^*A) = \lambda(XB^*Y^*YBX^*) = \lambda(B^*B)$$

and hence

$$E_r[\lambda(B^*B)] = E_r]\lambda(D^2B^*G^2B)]$$

for all B. Choose B as follows:

$$B = \left[\begin{array}{ccc|c} 0 & & 1 & \\ & \cdot & & \\ & \cdot & & 0 \\ & \cdot & & \\ 1 & & 0 & \\ \hline & 0 & & 0 \end{array}\right]$$

in which the upper left block is the indicated r-square permutation matrix. Then clearly $E_r[\lambda(B^*B)] = 1$ and

$$D^2B^*G^2B = \left[\begin{array}{ccccc|c} d_1^2g_r^2 & & & & & \\ & d_2^2g_{r-1}^2 & & 0 & & \\ & & \ddots & & & 0 \\ & 0 & & d_{r-1}^2g_2^2 & & \\ & & & & d_r^2g_1^2 & \\ \hline & & 0 & & & 0 \end{array}\right]$$

Thus

$$1 = E_r[\lambda(B^*B)] = \prod_{j=1}^{r} d_j^2 g_j^2 \ .$$

Now set $D^2 = RD_\sigma^2R$ where R is an n-square permutation matrix and D_σ^2 is a diagonal matrix obtained from D^2 by a permutation σ of the diagonal elements of D^2. Then

$$\begin{aligned} \lambda(D^2B^*G^2B) &= \lambda(RD_\sigma^2R^*B^*G^2B) \\ &= \lambda(D_\sigma^2(BR)^*G^2(BR)) \\ &= \lambda(D_\sigma^2C^*G^2C) \ , \end{aligned}$$

where $C = BR$, and

$$\lambda(B^*B) = \lambda(R^*B^*BR) = \lambda(C^*C) \ .$$

Therefore

$$E_r[\lambda(C^*C)] = E_r[\lambda(D_\sigma^2C^*G^2C)]$$

for all C. It follows that

$$\prod_{i=1}^{r} d_{\sigma(i)}^2 g_i^2 = 1$$

for any permutation σ of $1, \cdots, n$. From this we conclude that

$$d_1^2 = \cdots = d_n^2 = d^2$$

and similarly

$$g_1^2 = \cdots = g_n^2 = g^2 \ .$$

Then $G = gI$, $D = dI$ and $U = gQ$, $V = dP$, i.e. U, V are scalar multiples of unitary matrices. Now,

$$\begin{aligned} E_r[\lambda(A^*A)] &= E_r[\lambda((UAV)^*(UAV))] \\ &= E_r[\lambda(|g|^2 V^*A^*AV)] \\ &= E_r[\lambda(|gd|^2 A^*A)] \\ &= |gd|^{2r} E_r[\lambda(A^*A)] \, . \end{aligned}$$

Hence $|gd|^{2r} = 1$ and we can choose U and V to be gdQ and P which are unitary. This completes the proof.

We remark that in case $r = 1$ T does not necessarily have the form indicated in (2) and (3). For

$$E_1[\lambda(A^*A)] = tr(A^*A) = \sum_{(i,j)=(1,1)}^{(m,n)} |a_{ij}|^2 \, ,$$

and if T is merely a unitary operator on $M_{m,n}$

$$E_1[\lambda((T(A))^* T(A))] = E_1[\lambda(A^*A)] \, .$$

For example T can be the operator that interchanges the (1, 2) and (2, 1) elements of every $A \in M_{m,n}$ (assume $m, n > 2$) and this cannot be accomplished by any pre- and post-multipliplication as in (2) or (3).

References

1. J. Dieudonné, *Sur une géneralization du groupe orthogonal à quatre variables*, Archiv. d. Math., **1** (1948), 282-287.
2. G. Frobenius, *Über die Darstellung der endlichen Gruppen durch lineare Substitutionen*, Sitzungsberichte der Berliner Akademie, pp. 994-1015.
3. L. K. Hua, *Geometries of matrices. I. Generalizations of von Staudt's theorem*; Trans. Amer. Math. Soc., **57** (1945), 441-481.
4. H. G. Jacob, *Coherence invariant mappings on Kronecker products*, Amer. J. Math., **77** (1955), 177-189.
5. S. Kantor, *Theorie der Äquivalenz von linearen* ∞^λ *Scharen bilinearer Formen*, Sitzungsberichte der Münchener Akademie, (1897), 367-381.
6. M. Marcus and R. Westwick, *Linear map on skew-symmetric matrices: the invariance of elementary symmetric functions*, Pacific J. Math., **10** (1960), 917-924.
7. M. Marcus and B. N. Moyls, *Transformations on tensor product spaces*, Pacific J. Math., **9** (1959). 1215-1221.
8. M. Marcus and R. Purves, *Linear transformations on algebras of matrices: the invariance of the elementary symmetric functions*, Canadian J. Math., **11** (1959), 383-396.
9. K. Morita, *Schwarz's lemma in a homogeneous space of higher dimensions*, Japan J. of Math., **19** (1944), 45-56.
10. I. Schur, *Einige Bemerkungen zur Determinantentheorie*, Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin, **25** (1925), 454-463.

U. S. National Bureau of Standards, Washington, D. C.

The University of Florida, Gainesville, Florida