SIMULTANEOUS UNITARY INVARIANTS FOR SETS OF MATRICES

HEYDAR RADJAVI

It is our aim in this paper to give an elementary solution to the problem of simultaneous unitary equivalence of two finite sets of matrices, i.e., given two ordered sets $\{A_j\}$ and $\{B_j\}$ of $n \times n$ matrices, $j = 1, 2, \ldots, m$, we wish to determine whether there exists a unitary matrix U such that $B_j = U^*A_jU$ for all j. A special case of this problem is that of unitary equivalence of two arbitrary matrices.

If the process of diagonalizing a hermitian matrix is counted as a single "step", the solution presented here would give, in a finite number of steps, a complete set of numerical unitary invariants for each finite set of matrices. This solution is perhaps more compact and more practical than the existing ones, e.g., canonical forms discussed in (**1**)–(**4**).

For the sake of brevity, we shall, in the remainder of the paper, refer to "simultaneous unitary equivalence" and also to "unitary equivalence" as mere "equivalence". It will be tacitly assumed throughout the paper that sets of matrices are ordered.

1. The problem of equivalence of two arbitrary sets $\{A_j\}$ and $\{B_j\}$, $j = 1, 2, \ldots, m$, of $n \times n$ matrices over the field of complex numbers immediately reduces to that of equivalence of sets of $n \times n$ hermitian matrices, because every matrix can be uniquely expressed as $M + iN$, where both M and N are hermitian, and two matrices are equivalent if and only if their corresponding hermitian "components" are equivalent as sets. Hence, if $A_j = M_j + iN_j$ and $B_j = P_j + iQ_j$, where the matrices M_j, N_j, P_j, and Q_j are hermitian, then the sets $\{A_1, A_2, \ldots, A_m\}$ and $\{B_1, B_2, \ldots, B_m\}$ are equivalent if and only if the new sets $\{M_1, M_2, \ldots, M_m, N_1, N_2, \ldots, N_m\}$ and $\{P_1, P_2, \ldots, P_m, Q_1, Q_2, \ldots, Q_m\}$ are equivalent.

Notation and definitions. If G is a subgroup of the group of unitary matrices of size n and if A and B are two matrices of size n such that $B = U^*AU$, where U is a member of G, then A and B are said to be *equivalent under* G. If the subgroup G consists of all matrices of the form $\mathrm{Diag}(U, V) = U \oplus V$, where U and V are unitary matrices of size p and $n - p$, respectively, with $1 \leqq p < n$, then we call G a *direct group* and denote it by $G_n(p)$; the entire

Received December 29, 1966. This paper is based on part of the work done under the supervision of Professor Gerhard K. Kalisch and supported by the National Science Foundation under grant G 14137.

group of unitary matrices of size n is denoted by $G_n(n)$. If a matrix A of size n is written in the form

$$\begin{Vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Vmatrix},$$

where A_{11} is a square block of size p, we say that A is *partitioned into blocks conforming with* $G_n(p)$. If m is a positive integer, then the identity matrix of size m will be designated by $I(m)$.

In the following theorem, the symbol X_{ij} stands for the (i, j) entry of the matrix X.

THEOREM 1. *Given a positive integer p and a hermitian matrix A of size n with $p \leqq n$, let $S = \{U^*AU: U \in G_n(p)\}$. Then there exist a positive integer q, $1 \leqq q \leqq p$, a real number r, a non-negative number c, and a member B_0 of S such that*

(i) *B_0 has the form*

$$\begin{Vmatrix} rI(q) & N \\ N^* & M \end{Vmatrix},$$

with $NN^* = cI(q)$,

(ii) $r = \max\{B_{11}: B \in S\}$,

(iii) $c + r^2 = \max\{(B^2)_{11}: B \in S, B_{11} = r\}$, *and*

(iv) *if*

$$\begin{Vmatrix} rI(q_1) & N_1 \\ N_1^* & M_1 \end{Vmatrix},$$

is any member of S with $N_1N_1^* = cI(q_1)$, then $q_1 \leqq q$.

We shall call the matrix B_0 a *maximal transform* of A *relative to* $G_n(p)$. The integer q will be called the *reduction degree of A relative to $G_n(p)$*.

The method used in the following proof permits us to construct the numbers r, s, and q, and the matrices M and N.

Proof. Partition A into blocks conforming with $G_n(p)$:

$$A = \begin{Vmatrix} K & C \\ C^* & L \end{Vmatrix}.$$

The transform of A under a typical member $W = U \oplus V$ of $G_n(p)$ will be

$$W^*AW = \begin{Vmatrix} U^*KU & U^*CV \\ V^*C^*V & V^*LV \end{Vmatrix}.$$

The element $(W^*AW)_{11} = (U^*KU)_{11}$ will be maximal if and only if it is the greatest eigenvalue of the hermitian matrix K. Let r denote this eigenvalue and p_1 its multiplicity. If U is chosen to be a unitary matrix U_0 such that U_0^*KU is a diagonal matrix with diagonal elements in descending order, and if $W_0 = U_0 \oplus I(n-p)$, then

$$U_0^*KU_0 = \begin{Vmatrix} rI(p_1) & 0 \\ 0 & D \end{Vmatrix},$$

https://doi.org/10.4153/CJM-1968-098-4 Published online by Cambridge University Press

and the matrix $A_1 = W_0^* A W_0$ can be written in the following partitioned form:

$$A_1 = \begin{Vmatrix} rI(p_1) & 0 & E \\ 0 & D & F \\ E^* & F^* & L \end{Vmatrix}.$$

Now let U_1 be a $p_1 \times p_1$ unitary matrix such that $U_1^* EE^* U_1$ is diagonal with eigenvalues in descending order. Let c be the greatest eigenvalue of EE^* and q its multiplicity. If $W_1 = U_1 \oplus I(n - p_1)$, then the partition of $B_0 = W_1^* A_1 W_1$ conforming with $G_n(q)$ will be

$$B_0 = \begin{Vmatrix} rI(q) & N \\ N^* & M \end{Vmatrix},$$

where $NN^* = cI(q)$. The matrix B_0 and the number r satisfy conditions (i) and (ii).

We now prove (iii) and (iv). Let $B \in S$ with $B_{11} = B_{22} = \ldots = B_{tt} = r$. In the partition of B conforming with $G_n(p)$, the (1, 1) block is of the form $\operatorname{Diag}(rI(t), R)$, because r is the greatest eigenvalue of this block. Hence, if U is a unitary matrix of size $p - t$ which diagonalizes R with eigenvalues in descending order, then $W = I(t) \oplus U \oplus I(n - p) \in G_n(p)$ and $W^* B W \in S$. Since the first t rows of $W^* B W$ are the same as those of B, we shall assume, without loss of generality, that B is itself of the form

$$B = \begin{Vmatrix} rI(p_1) & 0 & E_0 \\ 0 & D_0 & F_0 \\ E_0^* & F_0^* & L_0 \end{Vmatrix}.$$

We observe that the matrix A_1 constructed above is equivalent to B under $G_n(p)$, and direct computation using the equation $WB = A_1 W$, $W \in G_n(p)$, shows, first, that W is necessarily of the form $U_1 \oplus U_2 \oplus U_3$, where U_1 is $p_1 \times p_1$ and U_2 is $(p - p_1) \times (p - p_1)$, and secondly, that $E_0 E_0^* = EE^*$. Since c is the greatest eigenvalue of EE^*, we have that

$$(B^2)_{jj} = r^2 + (E_0 E_0^*)_{jj} \leqq r^2 + c, \qquad j = 1, 2, \ldots, t.$$

If $t = 1$, the above argument proves (iii). Since the eigenvalue c of EE^* has multiplicity q, we conclude that if $(B^2)_{jj} = r^2 + c$ for $j = 1, 2, \ldots, t$, then $t \leqq q$, which implies (iv).

COROLLARY 1. *Let A be a matrix of size n and of the form*

$$\begin{Vmatrix} rI(p) & N \\ N^* & M \end{Vmatrix},$$

where $NN^* = cI(p)$. If q is any positive integer $\leqq p$, then any maximal transform of A relative to $G_n(q)$ is of the form

$$\begin{Vmatrix} rI(q) & Q \\ Q^* & P \end{Vmatrix}$$

with $QQ^* = cI(q)$.

COROLLARY 2. *Let q be the reduction degree of A relative to $G_n(p)$. If A_1 and A_2 are any two maximal transforms of A relative to $G_n(p)$, then they are equivalent under $G_n(q)$.*

COROLLARY 3. *If A and B are equivalent under $G_n(p)$, then they give rise to the same reduction degree q and their respective maximal transforms relative to $G_n(p)$ are equivalent under $G_n(q)$. This implies that the maximal transforms of A and B have the same r and the same c.*

2. We now present a lemma, a definition, and two theorems which will lead to the main result of this paper, namely, Theorem 4.

LEMMA 1. *Let A and B be two $p \times q$ matrices. Then the equation $B^*B = A^*A$ implies the existence of a unitary matrix U of size p such that $B = UA$.*

Proof. (i) If $p = q$, the assertion follows from the well-known theorem that every square matrix equals the product of a unitary matrix and a non-negative hermitian matrix. (ii) If $p > q$, we consider the square matrices $||A, 0||$ and $||B, 0||$, where 0 is the $p \times (p - q)$ zero matrix. It follows from (i) that $||B, 0|| = U||A, 0|| = ||UA, 0||$, and therefore $B = UA$ for some unitary matrix U. (iii) If $p < q$, the rank of the matrix $A^*A = B^*B$ is at most p, so that A^*A can be transformed, by some unitary matrix V, into the form $V^*A^*AV = V^*B^*BV = \mathrm{Diag}(M, 0)$, where M is a square block of size p. It follows that if we write $AV = ||A_1, A_2||$ and $BV = ||B_1, B_2||$, where A_1 and B_1 are square blocks of size p, then $A_2 = B_2 = 0$ and $A_1^*A_1 = M = B_1^*B_1$. We conclude, applying (i) again, that $B_1 = UA_1$ for some unitary matrix U, or $BV = ||UA_1, 0|| = U||A_1, 0|| = UAV$, which implies $B = UA$.

Definition. Let $\{A_j\}$, $j = 1, 2, \ldots, m$, be a set of hermitian matrices of size n. Let $U_1^*A_1U_1$ be a maximal transform of A_1 relative to $G_n(n)$ and let q_1 be the corresponding reduction degree. Having defined U_j and q_j for $1 \leqq j < k$, let

$$U_k^* U_{k-1}^* \ldots U_2^* U_1^* A_k U_1 U_2 \ldots U_{k-1} U_k$$

be a maximal transform of

$$U_{k-1}^* \ldots U_2^* U_1^* A_k U_1 U_2 \ldots U_{k-1}$$

relative to $G_n(q_{k-1})$ and let q_k be the corresponding reduction degree. The set of hermitian matrices $\{A'_j\}$, where

$$A'_j = U_m^* U_{m-1}^* \ldots U_2^* U_1^* A_j U_1 U_2 \ldots U_{m-1} U_m,$$

will be called a *reduced form* of the set $\{A_j\}$ and q_m the *reduction degree* of the set $\{A_j\}$.

THEOREM 2. *Let $\{A'_j\}$ be a reduced form of the set $\{A_j\}$ of hermitian matrices of size n and let q be the corresponding reduction degree with $q < n$. Then each A'_j has the form*

$$\begin{Vmatrix} r_j I(q) & N_j \\ N_j^* & M_j \end{Vmatrix},$$

where $N_j N_j^* = c_j I(q)$, $c_j \geqq 0$. *Furthermore,* $\{A_j\}$ *and* $\{B_j\}$ *are equivalent if and only if they have the same reduction degree* q *and their respective reduced forms* $\{A'_j\}$ *and* $\{B'_j\}$ *are equivalent under* $G_n(q)$.

Proof. The equivalence of $\{A_j\}$ and $\{B_j\}$ is, of course, implied by that of $\{A'_j\}$ and $\{B'_j\}$ under $G_n(q)$. The converse is proved by successive applications of Corollary 3 of Theorem 1.

Before stating Theorem 3 we make the following remarks.

(i) The reduction degree q of $\{A_j\}$ equals n if and only if each A_j is some multiple of $I(n)$.

(ii) We have $c_1 = 0$ by the definition of reduced forms.

(iii) If in the above theorem

$$A'_j = \begin{Vmatrix} r_j I(q) & N_j \\ N_j^* & M_j \end{Vmatrix} \quad \text{with } N_j N_j^* = c_j I(q)$$

and

$$B'_j = \begin{Vmatrix} s_j I(q) & Q_j \\ Q_j^* & P_j \end{Vmatrix} \quad \text{with } Q_j Q_j^* = d_j I(q),$$

then the equivalence of $\{A'_j\}$ and $\{B'_j\}$ under $G_n(q)$ implies $r_j = s_j$ and $c_j = d_j$ for all j.

THEOREM 3. *Let* $\{A_1, A_2, \ldots, A_m\}$ *and* $\{B_1, B_2, \ldots, B_m\}$ *be two sets of hermitian matrices of size* n, *where*

$$A_j = \begin{Vmatrix} r_j I(q) & N_j \\ N_j^* & M_j \end{Vmatrix} \quad \text{and} \quad B_j = \begin{Vmatrix} r_j I(q) & Q_j \\ Q_j^* & P_j \end{Vmatrix}$$

with $N_j N_j^* = Q_j Q_j^* = c_j I(q)$. *Assume* $c_k \neq 0$ *for some integer* k. *Then the two sets are equivalent under* $G_n(q)$ *if and only if the two sets*

$$\{M_1, M_2, \ldots, M_m, N_k^* N_1, N_k^* N_2, \ldots, N_k^* N_m\}$$

and

$$\{P_1, P_2, \ldots, P_m, Q_k^* Q_1, Q_k^* Q_2, \ldots, Q_k^* Q_m\}$$

of matrices of size $n - q$ *are equivalent.*

Proof. The equivalence of the new sets easily follows from that of the original sets under $G_n(q)$. We now prove the converse. The hypotheses imply that there exists a unitary matrix V of size $n - q$ such that $P_j = V^* M_j V$ and $Q_k^* Q_j = V^* N_k^* N_j V$, $j = 1, 2, \ldots, m$. Computation shows that for each j,

$$Q_j^* Q_k Q_k^* Q_j = V^* N_j^* N_k N_k^* N_j V.$$

But $Q_k Q_k^* = N_k N_k^* = c_k I(q) \neq 0$, so that $Q_j^* Q_j = V^* N_j^* N_j V$, $j = 1, 2, \ldots, m$. It follows from Lemma 1 that there exist m unitary matrices U_j

of size q such that $Q_j = U_j N_j V$, $j = 1, 2, \ldots, m$. Multiplying both members of the relation $Q_k^* Q_j = V^* N_k^* N_j V$ by the corresponding members of $Q_k = U_k N_k V$ on the left, and replacing Q_j by $U_j N_j V$ we obtain $U_j N_j = U_k N_j$. It follows that for each j with $c_j \neq 0$ we have $U_j N_j N_j^* = U_k N_j N_j^*$ or $U_j = U_k$. For each j with $c_j = 0$ the equality $Q_j = N_j = 0$ holds; hence, the matrix U_j can be chosen arbitrarily and we can put $U_j = U_k$. Then we have, for all j, $P_j = V^* M_j V$ and $Q_j = U_k N_j V$. This shows that the original sets $\{A_j\}$ and $\{B_j\}$ are equivalent under $G_n(q)$: $B_j = W^* A_j W$, $j = 1, 2, \ldots, m$, where $W = U_k^* \oplus V$. The proof is thus completed.

If in the statement of the above theorem we assume that $c_j = 0$ for all j, then $N_j = Q_j = 0$ for all j. In this case, $\{A_j\}$ and $\{B_j\}$ are equivalent under $G_n(q)$ if and only if $\{M_j\}$ and $\{P_j\}$ are equivalent under $G_{n-q}(n - q)$.

3. The matrices $N_k^* N_j$ and $Q_k^* Q_j$ in Theorem 3 are not, in general, hermitian; to replace the sets $\{M_j, N_k^* N_j\}$ and $\{P_j, Q_k^* Q_j\}$ by hermitian sets we split each of the matrices $N_k^* N_j$ and $Q_k^* Q_j$ into its hermitian components. For convenience, we adopt the notation

$$\mathrm{Re}(L) = (L + L^*)/2 \quad \text{and} \quad \mathrm{Im}(L) = (L - L^*)/2i.$$

Combining Theorems 2 and 3 with the preceding remarks, we obtain the following theorem.

THEOREM 4. *Let $\{A_i\}$ and $\{B_i\}$ be two sets of m hermitian matrices of size n. Let $\{A'_i\}$ and $\{B'_i\}$ be their respective reduced forms and p and q their respective reduction degrees. If*

$$A'_i = \begin{Vmatrix} r_i I(p) & N_i \\ N_i^* & M_i \end{Vmatrix} \quad \textit{and} \quad B'_i = \begin{Vmatrix} s_i I(q) & Q_i \\ Q_i^* & P_i \end{Vmatrix},$$

then $\{A_i\}$ and $\{B_i\}$ are equivalent if and only if $p = q$, $r_i = s_i$, $i = 1, 2, \ldots, m$, and the hermitian sets $\{M_i, \mathrm{Re}(N_0^* N_j), \mathrm{Im}(N_0^* N_j)\}$ and $\{P_i, \mathrm{Re}(Q_0^* Q_j), \mathrm{Im}(Q_0^* Q_j)\}$ are equivalent, $i = 1, 2, \ldots, m$, $j = 2, 3, \ldots, m$. The matrix N_0 is the first non-zero member of the set $\{N_2, N_3, \ldots, N_{m-1}\}$ if such a member exists; otherwise, $N_0 = N_m$.

This theorem reduces the problem of equivalence of two sets of m hermitian matrices of size n to that of equivalence of two new sets of hermitian matrices of size at most $n - 1$, where each set contains $3m - 2$ matrices. We use Theorem 4 to obtain the following.

Complete set of unitary invariants. Let $\{A_1, A_2, \ldots, A_m\}$ be a set of hermitian matrices of size n and $\{A'_1, A'_2, \ldots, A'_m\}$ one of its reduced forms. Let p_1 be the reduction degree of $\{A_i\}$ and let r_{1i} be the element lying in the first row and first column of A'_i. If $p_1 = n$, then the positive integer p_1 together with the m real numbers r_{1i} constitute a complete set of invariants. If $p_1 < n$, then each A'_i is of the form

https://doi.org/10.4153/CJM-1968-098-4 Published online by Cambridge University Press

$$\begin{Vmatrix} r_i I(p_1) & N_i \\ N_i^* & M_i \end{Vmatrix}.$$

We now consider the new set of $3m - 2$ hermitian matrices of size $n - p_1$:

$$\{M_1, \ldots, M_m; \operatorname{Re}(N_0^*N_2), \ldots, \operatorname{Re}(N_0^*N_m); \operatorname{Im}(N_0^*N_2), \ldots, \operatorname{Im}(N_0^*N_m)\},$$

where N_0 is defined as in Theorem 4. We repeat the above process, i.e., find the reduction degree p_2 and a reduced form of the new set and let r_{2i} be the (1, 1) element of the ith member of this reduced set, $i = 1, 2, \ldots, 3m-2$. If $p_1 + p_2 < n$, we keep repeating the process. Since each p_j is positive, after a finite number, k, of steps we will have $p_1 + p_2 + \ldots + p_k = n$. Then the following set of numbers forms a complete set of unitary invariants for $\{A_i\}$:

$$\begin{array}{l} p_1, p_2, \ldots, p_k; \\ r_{11}, r_{12}, \ldots, r_{1f(1)}; \\ r_{21}, r_{22}, \ldots, r_{2f(2)}; \\ \cdot \\ \cdot \\ \cdot \\ r_{k1}, r_{k2}, \ldots, r_{kf(k)}; \end{array}$$

where $f(j) = 3^{j-1}(m - 1) + 1$, the p_i are positive integers adding up to n, and the r_{ij} are real numbers.

In the special case where $\{A_i\}$ is a commutative set, i.e., $A_iA_j = A_jA_i$ for all i and j, we have $r_{kj} = 0$ for $j > m$.

In finding unitary invariants for an *arbitrary single matrix* A of size n, we consider the hermitian set $\{\operatorname{Re}(A), \operatorname{Im}(A)\}$, thus reducing the problem to the general case given above. It is interesting to note that if A is normal, then the hermitian set is commutative; hence the set of invariants becomes

$$\begin{array}{l} p_1, p_2, \ldots, p_k; \\ r_{11}, r_{12}; \\ r_{21}, r_{22}, 0, 0; \\ r_{31}, r_{32}, 0, 0, \ldots, 0; \\ \cdot \\ \cdot \\ \cdot \\ r_{k1}, r_{k2}, 0, 0, \ldots, 0; \end{array}$$

where the numbers $r_{j1} + ir_{j2}$ are the distinct eigenvalues of the matrix A and the p_j their corresponding multiplicities.

In conclusion, we note that the results of this paper apply not only to finite sets, but also to finitely generated infinite groups or rings of matrices.

References

1. J. Brenner, *The problem of unitary equivalence*, Acta Math. *86* (1951), 297–308.
2. D. E. Littlewood, *On unitary equivalence*, J. London Math. Soc. *28* (1953), 314–322.
3. B. E. Mitchell, *Unitary transformations*, Can. J. Math. *6* (1954), 69–72.
4. Heydar Radjavi, *On unitary equivalence of arbitrary matrices*, Trans. Amer. Math. Soc. *104* (1962), 363–373.

University of Minnesota,
Minneapolis, Minnesota;
University of Toronto,
Toronto, Ontario