Principal Minors of Complex Symmetric and Skew Matrices*

R. C. Thompson
Institute for Applied Algebra
Mathematics Department
University of California
Santa Barbara, California 93106

To Alston Householder
Submitted by Emeric Deutsch

ABSTRACT

The singular values of principal submatrices of complex symmetric and skew matrices are examined.

The singular values of a real or complex matrix are the eigenvalues of the Hermitian factor in its polar decomposition. These real and nonnegative quantities are of considerable interest in both numerical linear algebra and functional analysis, through the connection with the theory of norms, and it is generally believed that their properties are well understood. Recent developments [6, 7], however, have shown that this is not the case, and in particular the results in [7] have uncovered quite unexpected properties of the singular values of complex symmetric matrices. The purpose of the present note is to show another curious property of the singular values of complex symmetric matrices. Corresponding results for complex skew matrices and Hermitian matrices will also be established.

Let $A$ be a fixed $n \times n$ complex matrix with singular values $s_1 \geq \cdots \geq s_n$. Set $B = UAV$, where $U$ and $V$ are unitary, and take $B_n$ to be the leading $(n-1) \times (n-1)$ principal submatrix of $B$. Let $\sigma_1 \geq \cdots \geq \sigma_{n-1}$ be the singular values of $B_n$. The singular value interlacing inequalities assert that

$$
\begin{align*}
  s_1 &\geq \sigma_1 \geq s_3, \\
  s_2 &\geq \sigma_2 \geq s_4, \\
  \vdots & \quad \vdots \quad \vdots \\
  s_{n-2} &\geq \sigma_{n-2} \geq s_n, \\
  s_{n-1} &\geq \sigma_{n-1} \geq 0.
\end{align*}
$$

(1)
These inequalities were discovered by Thompson [5], the partial result $a_i < s_i$ $(1 < i < n)$ having been noticed earlier by Mirsky [2]. The converse result, asserting the completeness of (1), is due to Thompson [5], and states that as $U, V$ range over all unitary matrices, the singular values of $B_n = (UAV)_n$ assume all values $a_{1} > \cdots > a_{n-1}$ satisfying (1). It is interesting to compare the singular value interlacing inequalities with the well-known interlacing inequalities [1] for the eigenvalues of $(n - 1) \times (n - 1)$ principal submatrices of $n \times n$ Hermitian matrices. The chief reason that the index on the lower side of each inequality (1) drops by two (instead of by one as in the usual interlacing inequalities) is that the number of degrees of freedom in the singular value case is doubled, since we consider $UAV$ for variable unitary $U, V$, in contrast to $UAU^*$ for variable unitary $U$ when $A$ is Hermitian and eigenvalues are being studied. ($U^*$ is the adjoint of $U$.)

Now let $A$ be a complex symmetric matrix. To preserve the symmetry of $B = UAV$ we shall take $V = U^\dagger$, the transpose of $U$. Then the singular values of the leading principal submatrix $B_n$ of $B = UAU^\dagger$ still satisfy (1), and we may ask: Is (1) still complete? That is, as $U$ ranges over all unitary matrices, will the singular values of $B_n$ assume all values permitted by (1)? In view of the remark in the last paragraph about degrees of freedom, one might expect this not to be the case. The following result therefore is somewhat unexpected.

**Theorem 1.** Let $A$ be a complex symmetric matrix with singular values $s_1 > \cdots > s_n$. Set $B = UAU^\dagger$, where $U$ is unitary, and take $B_n$ to be the leading principal $(n - 1) \times (n - 1)$ submatrix of $B$. Then: as $U$ ranges over all unitary matrices, the singular values $a_{1} > \cdots > a_{n-1}$ of $B_n$ fill out all values permitted by the inequalities (1).

Next, let $A$ be a complex skew symmetric matrix. It is then known [8] that the nontrivial singular values occur in pairs; after changing notation we may take the singular values of $A$ to be

$$s_1, s_1, s_2, s_2, \ldots, s_m, s_m \quad \text{when } n = 2m,$$

$$s_1, s_1, s_2, s_2, \ldots, s_m, s_m, 0 \quad \text{when } n = 2m + 1,$$

where $s_1 > \cdots > s_m$. With $B = UAU^\dagger$ as before, the leading principal submatrix $B_n$ is again skew; let its singular values be

$$\sigma_1, \sigma_1, \sigma_2, \sigma_2, \ldots, \sigma_{m-1}, \sigma_{m-1}, 0 \quad \text{when } n = 2m,$$

$$\sigma_1, \sigma_1, \sigma_2, \sigma_2, \ldots, \sigma_m, \sigma_m \quad \text{when } n = 2m + 1.$$
Then the inequalities (1) take the form

\[
\begin{align*}
&\sigma_1 > \sigma_2 > \cdots > \sigma_{m-1} > \sigma_m > s_m & \text{ when } n = 2m, \\
&\sigma_1 > \sigma_2 > \cdots > \sigma_{m-1} > \sigma_m > s_m > 0 & \text{ when } n = 2m + 1.
\end{align*}
\]

(2)

**Theorem 2.** Let \( A \) be a complex skew matrix, and \( B = UAU^\dagger \) as in Theorem 1. Then, as \( U \) ranges over all unitary matrices, the singular values of the leading principal submatrix \( B_n \) of \( B \) fill out all values permitted by the inequalities (2).

We next state that Theorem 1 is not valid for real symmetric matrices, i.e., when \( A \) is real symmetric and \( U \) is real orthogonal. In contrast to this, Theorem 2 is valid for real skew matrices (\( A \) real skew, \( U \) real orthogonal). The real case analogous to Theorem 1 is the following.

**Theorem 3.** Let \( A \) be a real symmetric matrix with singular values \( s_1 > \cdots > s_n \), so that the eigenvalues of \( A \) are \( \lambda_1 = \pm s_1, \ldots, \lambda_n = \pm s_n \) for a certain choice of the \( \pm \) signs. Then, as \( U \) ranges over all real orthogonal matrices, the singular values of \( B_n \) (where \( B = UAU^\dagger \)) fill out all values \( \sigma_1 > \cdots > \sigma_{n-1} \) allowed by (1) if and only if the following conditions are satisfied:

(a) \( \sigma_i = \sigma_i^1 \) and \( \sigma_{i+1} = \sigma_{i+1}^1 \) carry opposite signs whenever \( s_i \) and \( s_{i+1} \) are both nonzero simple singular values;

(b) each nonzero multiple singular value has at least one eigenvalue of each sign.

Theorem 3 remains valid when \( A \) is Hermitian and \( B = UAU^* \).

**Proof of Theorem 1.** We first note that \( A = WDW^\dagger \) for some unitary \( W \), where \( D = \text{diag}(s_1, s_2, \ldots, s_n) \). This decomposition of a complex symmetric matrix, not difficult to prove, is due to Schur [3]. (Or see [8].) Thus no generality is lost if we take \( A = \text{diag}(s_1, s_2, \ldots, s_n) \).

We now attach signs to the \( s_i \) and \( \sigma_i \) in this manner:

\[
\begin{align*}
& s_1, -s_2, s_3, -s_4, s_5, -s_6, \text{ etc.,} & (3) \\
& \sigma_1, -\sigma_2, \sigma_3, -\sigma_4, \sigma_5, -\sigma_6, \text{ etc.} & (4)
\end{align*}
\]

Our assumptions then imply that the numbers (4) interlace the numbers (3) when these two sets of numbers are rearranged into nonincreasing order.
Consequently [1], a real orthogonal matrix $\Theta$ exists such that the leading $(n-1)$-square submatrix in

$$
\Theta \text{ diag}(s_1, -s_2, s_3, -s_4, \ldots) \Theta^T
$$

has eigenvalues (4). Set $D = \text{diag}(1, i, 1, i, \ldots)$, where $i = (-i)^{1/2}$; then put $U = \Theta D$. This matrix $U$ is unitary; thus $B = U \text{diag}(s_1, \ldots, s_n) U^T$ has $s_1, \ldots, s_n$ as its singular values, and the choice of $\Theta$ makes $B_n$ real symmetric with eigenvalues $\pm \sigma_i$, i.e., with singular values $\sigma_1, \ldots, \sigma_{n-1}$.

Proof of Theorem 2. A known result [8] asserts that any $n \times n$ complex skew matrix has the form $WDW^\dagger$, where $W$ is unitary and $D$ is block diagonal with $2 \times 2$ diagonal blocks of the form

$$
\begin{bmatrix}
0 & s \\
-s & 0
\end{bmatrix},
$$

$s \geq 0,$

therefore suffice to assume that $A = D$ is block diagonal of this type. We wish to find a real orthogonal $\Theta$ such that $(\Theta D \Theta^\dagger)_{n}$ has the prescribed singular values. We write out the following proof in the odd dimensional case, the even dimensional case differing only in minor ways.

Let

$$
D = \begin{bmatrix}
0 & s_1 \\
-s_1 & 0
\end{bmatrix} + \cdots + \begin{bmatrix}
0 & s_m \\
-s_m & 0
\end{bmatrix} + [0].
$$

Take $B = \Theta D \Theta^\dagger$. Then $\lambda I - B = \Theta (\lambda I - D) \Theta^\dagger$, where $\lambda$ is a polynomial indeterminate. Taking inverses and multiplying by $\det(\lambda I - B)$, we get

$$
\text{adj}(\lambda I - B) = \Theta \{\text{adj}(\lambda I - D)\} \Theta^\dagger,
$$

adj denoting adjugate. Reading off the $(n, n)$ entry on each side, we then have ($\langle \; \rangle$ denoting a deleted factor)

$$
f(\lambda) = \sum_{i=1}^{m} (\lambda^2 + s_i^2) \cdots (\lambda^2 + s_i^2) \cdots (\lambda^2 + s_m^2) \lambda^2 (x_i^2 + y_i^2)
$$

$$
+ (\lambda^2 + s_i^2) \cdots (\lambda^2 + s_m^2) z^2,
$$

(5)

where $(x_1, y_1, \ldots, x_m, y_m, z)$ is the last row of $\Theta$, and $f(\lambda)$ is the characteristic
polynomial of $B_m$. Assume for the moment that $s_1, \ldots, s_m$ are distinct and nonzero.

We now choose special real values $x_1, y_1, \ldots, x_m, y_m, z$ such that the polynomial $f(\lambda)$ defined by (5) takes a particular form. From this particular form it will follow that $\Sigma_i (x_i^2 + y_i^2) + z^2 = 1$, and thus that our $x_i, y_i, z$ can be the last row of some real orthogonal matrix $\Theta$. Choose real $x_i, y_i, z$ to satisfy
\[
x_i^2 = y_i^2 = -\frac{\prod_{j=1}^{m} (s_j^2 - s_i^2)}{2s_i^2 \prod_{j=1}^{m} (s_j^2 - s_i^2)}, \quad 1 \leq i \leq m,
\]
\[
z = \sigma_1 \cdots \sigma_m / s_1 \cdots s_m.
\]
The conditions (2) ensure that real numbers $x_i, y_i, z$ exist satisfying these conditions.

This choice of $x_i, y_i, z$ forces $f(\lambda)$ to assume the same value as the polynomial $g(\lambda) = (\lambda^2 + \sigma_1^2) \cdots (\lambda^2 + \sigma_m^2)$ when any of $\pm is_1, \ldots, \pm is_m, 0$ are substituted for $\lambda$. Thus the degree $2m$ polynomials $f(\lambda)$ and $g(\lambda)$ agree at $2m+1$ distinct points and hence are the same polynomial. Comparing leading coefficients, we get $\Sigma_i (x_i^2 + y_i^2) + z^2 = 1$, as desired.

It now follows that a real orthogonal matrix $\Theta$ exists with these numbers $x_1, y_1, \ldots, x_m, y_m, z$ as its last row. For this $\Theta$, with $B = \Theta D \Theta^\dagger$, we find that the characteristic polynomial $f(\lambda)$ of $B_n$ equals $g(\lambda)$, so that $B_n$ has eigenvalues $\pm i\sigma_1, \ldots, \pm i\sigma_m$, and hence has the required singular values.

The existence of $B$ in case of nondistinct and possibly zero $s_i$ then follows by a standard limiting argument, passing to this case by continuity.

Proof of Theorem 3. Let $A$ be Hermitian, with eigenvalues $\pm s_1, \ldots, \pm s_n$. Suppose that $s_{i+1}$ and $s_{i+2}$ are both simple singular values, and that both occur as positive eigenvalues of $A$ or both occur as negative eigenvalues of $A$. The following numbers $\sigma_1, \ldots, \sigma_{n-1}$ are consistent with (1): $\sigma_1 = s_1, \ldots, \sigma_i = s_i, \sigma_{i+1} = s_{i+3}, \ldots, \sigma_{n-2} = s_n, \sigma_{n-1} = 0$. The eigenvalues $\beta_1, \ldots, \beta_{n-1}$ of $B_n$ (where $B = UAU^*$, $U$ unitary) must interlace $\lambda_1 = \pm s_1, \ldots, \lambda_n = \pm s_n$. Since $s_{i+1}$ and $s_{i+2}$ have the same sign, they are consecutive eigenvalues of $A$. The interval with endpoints $s_{i+1}, s_{i+2}$ must therefore contain at least one of $\pm s_{i-1}, \ldots, \pm s_{n-1}$; however, it does not, because $\pm s_1, \ldots, \pm s_i, \pm s_{i+3}, \ldots, \pm s_n, 0$ all lie outside this interval. Condition (a) of Theorem 3 therefore is necessary.
As for condition (b), suppose that \( s_i > s_{i+1} = \cdots = s_{i+e} > s_{i+e+1} \) with multiplicity \( e > 1 \), and that the corresponding eigenvalues are all of one sign, say positive. Take \( \sigma_1, \ldots, \sigma_{n-1} \) as in the last paragraph. We are required to assign \( \pm \) signs to the \( \sigma_i \) so that \( \pm \sigma_1, \ldots, \pm \sigma_{n-1} \) interlace \( \lambda_1, \ldots, \lambda_t \). Thus we have to take \( e-1 \) of the \( \pm \sigma_1, \ldots, \pm \sigma_{n-1} \) equal to \( s_{i+1} \). However, only \( e-2 \) of the \( \sigma_{i+1} \) equal \( s_{i+1} \), so this is not possible. Condition (b) therefore is necessary.

We now suppose (1) holds, with conditions (a) and (b) satisfied.

As a first case, suppose that each singular value \( s_i \) is simple. Then the eigenvalues \( \lambda_i \) are \( s_1, -s_2, s_3, -s_4, \) etc., or the negative of these, and by setting \( B = \emptyset \text{ diag}(s_1, -s_2, s_3, -s_4, \ldots) \emptyset^t \) as in the proof of Theorem 1, we find that \( B_n \) has eigenvalues \( \sigma_1, -\sigma_2, \sigma_3, -\sigma_4, \ldots \), and therefore the required singular values. So assume that multiple singular values exist. Our conditions ensure that if an \( s_i \) (for fixed \( i \)) occurs with multiplicity \( e > 1 \) among \( s_1, \ldots, s_n \), then \( s_i \) occurs with multiplicity \( e-2 \) (at least) among the \( \sigma_i \). Delete \( e-2 \) occurrences of \( s_i \) from \( s_1, \ldots, s_n \), by deleting \( e-2 \) \( \lambda_i \) satisfying \( |\lambda_i| = s_i \), retaining just two \( \lambda_i \) satisfying \( |\lambda_i| = s_i \), specifically retaining both \( s_i \) and \( -s_i \). This is possible by assumption (b). Also delete \( e-2 \) occurrences of \( s_i \) among \( \sigma_1, \ldots, \sigma_{n-1} \). Do these deletions for each multiple singular value. Let the remaining \( s_i, \sigma_i \), and \( \lambda_i \) be \( S_1 \geq \cdots \geq S_k; \Sigma_1 \geq \cdots \geq \Sigma_{k-1} \); and \( \Lambda_1 = \pm S_1, \ldots, \Lambda_k = \pm S_k \). Each \( S_i \) is then simple or has multiplicity precisely two.

We now attach \( \pm \) signs to the \( \Sigma_i \), as follows. If \( S_i = S_{i+1} > S_{i+2} > \cdots > S_u = S_{u+1} \), attach signs in this manner: \( \Sigma_{t'}, -\Sigma_{t+1}, \Sigma_{t+2}, -\Sigma_{t+3}, \ldots, (-1)^{t-u-1}\Sigma_{u-1} \), if \( \Lambda_{t+1} > 0 \); otherwise attach signs in this way: \( -\Sigma_{t'}, \Sigma_{t+1}, -\Sigma_{t+2}, \Sigma_{t+3}, \ldots, (-1)^{t-u}\Sigma_{u-1} \). Further, if \( S_i \) is the largest multiple singular value and \( t > 1 \), we attach signs according to: \( \Sigma_{t'}, -\Sigma_{t+1}, \Sigma_{t+2}, -\Sigma_{t+3}, \ldots, (-1)^{t-1}\Sigma_{t-1} \), when \( \Lambda_i > 0 \), and the negatives of these when \( \Lambda_i < 0 \). And if \( S_i \) is the smallest multiple singular value and \( t < n \), we take \( \Sigma_{t'}, -\Sigma_{t+1}, \Sigma_{t+2}, -\Sigma_{t+3}, \) etc., unless \( t+2 < n \) and \( \Lambda_{t+2} < 0 \), in which case we take the negatives of these.

Then the signs assigned to \( \Sigma_1, \ldots, \Sigma_{k-1} \) are such that \( \pm \Sigma_1, \ldots, \pm \Sigma_{k-1} \) interlace \( \Lambda_1, \ldots, \Lambda_k \). We may therefore find a real orthogonal matrix \( \emptyset \) such that the leading deficiency one submatrix in \( \emptyset \text{ diag}(\Lambda_1, \ldots, \Lambda_n) \emptyset^t \) has eigenvalues \( \pm \Sigma_1, \ldots, \pm \Sigma_{k-1} \). Attaching \( \text{ diag}(\Lambda_1, \ldots, \Lambda_k) \) to a diagonal matrix of the \( \lambda_i \) deleted above (as a direct summand), attaching \( \emptyset \) to an identity matrix, and naming the resulting matrices \( A \) and \( \emptyset \), respectively, it now follows that \( B = \emptyset A \emptyset^t \) is a real symmetrix matrix with singular values \( s_1, \ldots, s_n \) such that the leading \( (n-1) \)-square submatrix \( B_n \) has singular values \( \sigma_1, \ldots, \sigma_{n-1} \). The proof is complete.

The preparation of this paper was supported in part by the U.S. Air Force Office of Scientific Research, under Grant 77-3166.
REFERENCES

5 R. C. Thompson, Principal submatrices IX: Interlacing inequalities for singular values of submatrices, Linear Algebra and Appl. 5:1–12 (1972).
7 R. C. Thompson, Singular values and diagonal elements of complex symmetric matrices, Linear Algebra and Appl., to appear.

Received 10 July 1978