Matrix Behaviour, Unitary Reducibility, and Hadamard Products *

D. Viswanath and L.N. Trefethen †

July 9, 1996

Abstract

The question investigated here is: if two matrices A and B in $C^{N,N}$ have identical behaviour in a unitarily invariant norm $\|\cdot\|$, i.e., $\|p(A)\| = \|p(B)\|$ for every polynomial p with complex coefficients, what properties do A and B have in common? After a preliminary result about eigenvalues, it is shown with a mildly restrictive assumption that if A is unitarily reducible, so is B. A theorem is proved about Hadamard products of the form $H \circ H^{-T}$, where H is Hermitian positive definite. Finally, an example is produced where A and B have identical behaviour in the Frobenius norm, but are not related to each other in any simple way.

By *matrix behaviour* of a square matrix A in a norm $\|\cdot\|$, we mean the variation of $\|p(A)\|$ as p ranges over all polynomials with complex coefficients [2]. Matrix behaviour provides vital information about the matrix. For example, for a discrete process defined by $x_{n+1} = Ax_n$, boundedness of $\|A^n\|$ is the condition for stability. Thus the task of understanding matrix behaviour is an important one. For a normal matrix, the eigenvalues determine behaviour in any unitarily invariant norm [3]. For a non-normal matrix, however, the eigenvalues do not determine behaviour. Eigenvalues have been generalized to pseudospectra with the aim of capturing the behaviour of non-normal matrices [5].

In this paper, we explore the converse question: if the behaviour of A in a unitarily invariant norm is known, what properties of A may be deduced?

*This work was supported by NSF Grant DMS-9500975CS and DOE Grant DE-FG02-94ER25199

†Department of Computer Science, Cornell University, Ithaca, NY 14853 (divakar@cs.cornell.edu and lnt@cs.cornell.edu)

Recall that a norm $\|\cdot\|$ on $C^{N,N}$ is *unitarily invariant* if $\|A\| = \|PAQ\|$ for every matrix A and any unitary matrices $P, Q \in C^{N,N}$.

Throughout this paper, we assume that N is a positive integer, that $\|\cdot\|$ is a unitarily invariant norm on $C^{N,N}$, the set of $N \times N$ matrices with complex entries, and that $A, B \in C^{N,N}$. We use $\mathrm{diag}(A_1, \ldots, A_n)$ to denote a block diagonal matrix with the square matrices A_i on the diagonal. When we write $\mathrm{diag}(A_1, \ldots, A_n)$, it is assumed that A_i are square matrices with dimensions summing to N and that the dimension of each A_i is at least 1.

1 Unitary Reducibility

The simple result below is stated for completeness.

Theorem 1.1. *Let $A, B \in C^{N,N}$ satisfy $\|p(A)\| = \|p(B)\|$ for every polynomial p with complex coefficients. Then the minimal polynomials of A and B are identical. In particular, A and B have the same eigenvalues, not counting multiplicities.*

Proof. By assumption, $\|p(A)\| = 0$ if and only if $\|p(B)\| = 0$. Therefore, $p(A) = 0$ exactly when $p(B) = 0$, implying that A and B have the same minimal polynomial. Since the eigenvalues of a matrix are the roots of its minimal polynomial, A and B have the same eigenvalues. $\square$

Definition. A matrix A is *unitarily reducible* if it is unitarily similar to a matrix $\mathrm{diag}(M, N)$. An equivalent statement is that there exist two complementary invariant subspaces of A that are mutually orthogonal, neither space being of dimension zero.

The main result of this section asserts that if $\|p(A)\| = \|p(\mathrm{diag}(S, T))\|$ for every polynomial p, where S and T are fixed matrices with no common eigenvalues, then A must be unitarily reducible. In other words, unitary reducibility may be determined from behaviour.

In the proof of the following lemma, $\sigma_i(A)$ denotes the ith largest singular value of A. In particular, $\sigma_1(A)$ is the 2-norm of A.

Lemma 1.2. *If*

$$\|\mathrm{diag}(I_{k_1}, 0)\| = \|V\,\mathrm{diag}(I_{k_2}, 0)V^{-1}\|,$$

where $k_1 \le k_2$ and $V \in C^{N,N}$ is nonsingular, then the first k_2 columns of V are orthogonal to the last $N - k_2$ columns.

Proof. Let $V = QR$ be a QR factorization of V, where

$$R = \begin{pmatrix} R_{11} & R_{12} \\ & R_{22} \end{pmatrix}, \quad R_{11} \in C^{k_2,k_2}.$$

It suffices to show that $R_{12} = 0$.

Transform the matrix $V\,\mathrm{diag}(I_{k_2}, 0)V^{-1}$ into a convenient form through a unitary similarity:

$$(Q^*V)\,\mathrm{diag}(I_{k_2}, 0)(V^{-1}Q) = R(\mathrm{diag}(I_{k_2}, 0))R^{-1} = \begin{pmatrix} I_{k_2} & -R_{12}R_{22}^{-1} \\ 0 & 0 \end{pmatrix}.$$

If

$$A = \begin{pmatrix} I_{k_1} & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} I_{k_2} & -R_{12}R_{22}^{-1} \\ 0 & 0 \end{pmatrix},$$

then $\|A\| = \|B\|$ by hypothesis.

If $R_{12} \neq 0$, B has a row with 2-norm greater than 1. Therefore $\sigma_1(B) > 1$. Since $k_1 \leq k_2$, the interlacing inequalities for singular values imply $\sigma_i(B) \geq \sigma_i(A)$ for every i [4]. Given that $\sigma_1(B) > \sigma_1(A) = 1$ and that $\sigma_i(B) \geq \sigma_i(A)$ for every i, the quasilinear characterization of unitarily invariant norms in Theorem 3.5.5 of [4] implies that $\|B\| > \|A\|$. This contradicts $\|A\| = \|B\|$. Therefore, $R_{12} = 0$ as desired.

□

Since the second half of the previous proof is interesting in its own right, we state it as a proposition.

Proposition 1.3. *Let*

$$A = \begin{pmatrix} I_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} I_2 & X \\ 0 & 0 \end{pmatrix},$$

where I_1 is the identity in C^{k_1,k_1}, I_2 is the identity in C^{k_2,k_2}, and $0 < k_1 \leq k_2 < N$. Then $\|A\| = \|B\|$ implies $X = 0$.

In the theorem below, the assumption that the minimal polynomials of S_1 and T_1 are prime to one another is equivalent to the assumption that S_1 and T_1 do not have any common eigenvalues. Recall that A and B are assumed to be matrices in $C^{N,N}$ throughout this paper.

Theorem 1.4. *Let $\|p(A)\| = \|p(B)\|$ for every polynomial p. Assume*

$$A = \operatorname{diag}(S, T), \quad S \in C^{k_1,k_1}, \quad T \in C^{N-k_1,N-k_1},$$

with the minimal polynomials of S_1 and T_1 prime to one another. Then B is unitarily similar to a matrix of the form

$$\operatorname{diag}(S', T'), \quad S' \in C^{k_2,k_2}, \quad T' \in C^{N-k_2,N-k_2}.$$

Moreover, for every polynomial p,

$$\|p(\operatorname{diag}(S, 0))\| = \|p(\operatorname{diag}(S', 0))\|,$$
$$\|p(\operatorname{diag}(0, T))\| = \|p(\operatorname{diag}(0, T'))\|.$$

Proof. Let s and t be the minimal polynomials of S and T respectively. Consider first the matrix A, st being its minimal polynomial. Since s and t are relatively prime, there exist polynomials g and h such that $sg + th = 1$. Since $t(S)h(S) = I - s(S)g(S) = I$ and $t(T)h(T) = 0$, and similarly $s(S)g(S) = 0$ and $s(T)g(T) = I$, we have

$$t(A)h(A) = \operatorname{diag}(I, 0), \qquad s(A)g(A) = \operatorname{diag}(0, I),$$

where the two identities have the same dimensions as S and T respectively. Thus $t(A)h(A)$ and $s(A)g(A)$ are complementary projectors onto the invariant subspaces of A associated with the eigenvalues of S and T respectively. These spaces are orthogonal, and thus these are orthogonal projectors.

Now consider the matrix B. Since $\|p(A)\| = \|p(B)\|$ for every p, B has the same minimal polynomial as A, namely st. Thus B is similar, even if not unitarily, to a block diagonal matrix whose blocks have the minimal polynomials s and t:

$$B = V \operatorname{diag}(S'', T'')V^{-1}, \quad S'' \in C^{k_2,k_2}.$$

For example, we can get this representation from the Jordan canonical form. As before, we calculate

$$t(B)h(B) = V \operatorname{diag}(I, 0)V^{-1}, \qquad s(B)g(B) = V \operatorname{diag}(0, I)V^{-1}.$$

Thus, as before, $t(B)h(B)$ and $s(B)g(B)$ are projections into complementary invariant subspaces of B.

Now we can use either $\|t(A)h(A)\| = \|t(B)h(B)\|$ or $\|s(A)g(A)\| = \|s(B)g(B)\|$, according as $k_1 \leq k_2$ or $N - k_1 \leq N - k_2$, with Lemma

1.2 to conclude that the first k_2 columns of V are orthogonal to the last $N - k_2$ columns. This means that the invariant subspaces associated with S'' and T'' are orthogonal. In other words, B is unitarily similar to a matrix $\mathrm{diag}(S', T')$ with s, t the minimal polynomials of S' and T'.

To prove the last part, let $B' = \mathrm{diag}(S', T')$. The matrix B' is unitarily similar to B. Since $xs(x)g(x) + xt(x)h(x) = x$, we have

$$At(A)h(A) = \mathrm{diag}(St(S)h(S), Tt(T)h(T)) = \mathrm{diag}(S, 0)$$

and

$$B't(B')h(B') = \mathrm{diag}(S't(S')h(S'), T't(T')h(T')) = \mathrm{diag}(S', 0).$$

By assumption, $\|p(At(A)h(A))\| = \|p(B't(B')h(B'))\|$ for every polynomial p, and hence $\|p(\mathrm{diag}(S, 0))\| = \|p(\mathrm{diag}(S', 0))\|$ for every polynomial p. Similarly, it can be shown that $\|p(\mathrm{diag}(0, T))\| = \|p(\mathrm{diag}(0, T'))\|$ for every polynomial p.

$\square$

The theorem above can be strengthened when $\|\cdot\| = \|\cdot\|_2$. In that case, we can say that $\|p(S)\|_2 = \|p(S')\|_2$ and $\|p(T)\|_2 = \|p(T')\|_2$ for every polynomial of p. Also, A and B do not have to have the same dimension.

2 A Result about Hadamard Products

The Hadamard product of $A = (a_{ij})$ and $B = (b_{ij})$ is their entry-wise product $A \circ B = (a_{ij} b_{ij})$. The next proposition, along with Theorem 1.4, suggests a new result about Hadamard products. It is also used in the next section.

Proposition 2.1. *Let A be diagonalizable with distinct eigenvalues. Let*

$$A = V\,\mathrm{diag}(d_1, \dots, d_n) V^{-1}, \quad d_i \in C,$$

be a diagonalization of A. If $p(x)$ is a polynomial with complex coefficients the Frobenius norm of $p(A)$ is given by

$$\|p(A)\|_F^2 = \rho^*(V^*V \circ (V^*V)^{-T})\rho,$$

where $\rho = (p(d_1), \dots, p(d_n))^T$.

Proof. See [1]. $\square$

Suppose $A = V\,\mathrm{diag}(d_1, \ldots, d_n)V^{-1}$ is unitarily reducible with the first k columns of V orthogonal to the last $N-k$ columns. Then V^*V is block diagonal, and hence $V^*V \circ (V^*V)^{-T}$ is also block diagonal. But does $V^*V \circ (V^*V)^{-T}$ being block diagonal imply that V^*V is block diagonal, and hence that A is unitarily reducible? The answer, as given by the next theorem, is yes.

The significance of Hadamard products of the form $H \circ H^{-T}$ and several of their properties are discussed in [4]. The matrix obtained by replacing every element of A by its complex conjugate is denoted by $\bar{A}$.

Lemma 2.2. *Let $y \in C^N$ and let $H \in C^{N,N}$ be Hermitian positive definite. If $y \circ H\bar{y} = 0$, then $y = 0$.*

Proof. Assume $y \neq 0$. Since $\bar{H}$ is also Hermitian positive definite, $y^*\bar{H}y > 0$. This implies that y and $\bar{H}y$ have a non-zero in the same component. Since $H\bar{y} = \overline{\bar{H}y}$, y and $H\bar{y}$ have a non-zero in the same component, and $y \circ H\bar{y} \neq 0$. This contradicts the assumption $y \circ H\bar{y} = 0$. $\square$

Lemma 2.3. *Let $H_1 \in C^{n,n}$ and $H_2 \in C^{k,k}$ be Hermitian positive definite. Let $Y \in C^{n,k}$. If $Y \circ H_1\bar{Y}H_2 = 0$, then $Y = 0$.*

Proof. Let Y' be the column vector of length nk obtained by writing down the columns of Y one below another. Then $Y \circ H_1\bar{Y}H_2 = 0$ is equivalent to $Y' \circ (H_2^T \otimes H_1)\bar{Y'} = 0$, where $\otimes$ denotes the Kronecker product. Since $(H_2^T \otimes H_1)$ is Hermitian positive definite [4], Lemma 2.2 implies $Y' = 0$. Hence $Y = 0$ as claimed. $\square$

A theorem due to Schur says that $A \circ B$ is Hermitian positive definite if A and B are Hermitian positive definite [4]. Therefore, if H is Hermitian positive definite, so is $H \circ H^{-T}$.

Theorem 2.4. *If $H \in C^{N,N}$ is Hermitian positive definite, the Hermitian positive definite matrix $H \circ H^{-T}$ is of the form*

$$\mathrm{diag}(S, T), \quad S \in C^{k,k}, \quad T \in C^{N-k,N-k}, \quad 0 < k < N, \tag{2.1}$$

if and only H is also of the same form.

Proof. If H is of the form (2.1), that $H \circ H^{-T}$ is also of the same form is trivial.

Conversely, assume that $H \circ H^{-T} = \mathrm{diag}(S, T)$ where $S \in C^{k,k}$. Let

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}, \quad H_{11} \in C^{k,k}, \quad H_{22} \in C^{N-k,N-k}.$$

By a well-known formula [3], we have

$$H^{-1} = \begin{pmatrix} C_{11}^{-1} & -H_{11}^{-1}H_{12}C_{22}^{-1} \\ -C_{22}^{-1}H_{12}^{*}H_{11}^{-1} & C_{22}^{-1} \end{pmatrix},$$

where $C_{11} = (H_{11} - H_{12}H_{22}^{-1}H_{12}^{*})$ and $C_{22} = (H_{22} - H_{12}^{*}H_{11}^{-1}H_{12})$. Clearly, C_{22}^{-1} is Hermitian positive definite since it is a principal submatrix of the Hermitian positive definite matrix H^{-1}.

By assumption, the upper-right $k \times (N-k)$ block of $H \circ H^{-T}$ is zero. Hence,

$$H_{12} \circ (H_{11}^{-T}\bar{H}_{12}C_{22}^{T}) = 0.$$

By Lemma 2.3, this implies that $H_{12} = 0$, i.e., H has the form (2.1).

□

3 An Example

This section presents an example that surprised us. We produce two unitarily irreducible matrices A and B with different singular values which, nevertheless, satisfy $\|p(A)\|_F = \|p(B)\|_F$ for every polynomial p. In other words, these two matrices behave identically in the Frobenius norm but have no obvious relationship to each other. The Frobenius norm of A, $\|A\|_F$, may be defined as the square root of either the sum of the squares of the absolute values of the entries of A or the sum of squares of the singular values of A.

The next lemma, an easily proved fact of linear algebra, considers the eigenspaces of two unitarily similar matrices.

Lemma 3.1. *Let $A, B \in C^{N,N}$ be unitarily similar, diagonalizable matrices with distinct eigenvalues. If $A = VDV^{-1}$ and $B = UDU^{-1}$ are diagonalizations of A and B, there exists a diagonal matrix D_{uv} with non-zero diagonal entries such that $V^*V = D_{uv}^*(U^*U)D_{uv}$.*

Proof. Let $A = QBQ^*$ where Q is unitary. Then $A = (QU)D(QU)^{-1}$ and $A = VDV^{-1}$ are two diagonalizations of A. The columns of QU and V are the eigenvectors of A. Since the eigenvalues of A are distinct, there is only one direction an eigenvector for a given eigenvalue can point. Therefore, there must be a diagonal matrix D_{uv} with non-zero diagonal entries such that $V = QUD_{uv}$. For such a D_{uv}, $V^*V = D_{uv}^*(U^*U)D_{uv}$. □

The idea that led to the example in the next theorem is explained after its proof.

Theorem 3.2. *There exist matrices $A, B \in C^{N,N}$ with $N > 3$ such that*

- $\|p(A)\|_F = \|p(B)\|_F$ *for all polynomials* p,
- *neither A nor B is unitarily reducible,*
- *A and B have different singular values,*
- *A is unitarily similar to neither B nor B^T.*

Proof. Take $N = 4$. Choose $U, V \in C^{4,4}$ so that:

$$U^*U = \begin{pmatrix} 9 & 0 & 0 & 3 \\ 0 & 2 & 1/4 & 1 \\ 0 & 1/4 & 2 & 1 \\ 3 & 1 & 1 & 5 \end{pmatrix},$$

$$V^*V = \begin{pmatrix} 1 & 0 & 0 & \sqrt{\frac{567}{1764}} \\ 0 & 1 & 0 & \sqrt{\frac{252}{1764}} \\ 0 & 0 & 1 & \sqrt{\frac{252}{1764}} \\ \sqrt{\frac{567}{1764}} & \sqrt{\frac{252}{1764}} & \sqrt{\frac{252}{1764}} & \frac{2835}{1764} \end{pmatrix}.$$

Both U^*U and V^*V are diagonally dominant and, hence, Hermitian positive definite. Choose $D = \text{diag}(1, 2, 3, 4)$; D can be any diagonal matrix with distinct diagonal entries. Consider the matrices

$$A = VDV^{-1}, \quad B = UDU^{-1}.$$

We will show that A and B have all the properties listed in the theorem.

It may be verified that $(V^*V) \circ (V^*V)^{-T} = (U^*U) \circ (U^*U)^{-T}$. Therefore, Proposition 2.1 implies that $\|p(A)\|_F = \|p(B)\|_F$ for all polynomials p.

A unitarily reducible matrix has complementary invariant subspaces that are orthogonal. Since A and B are both diagonalizable with distinct eigenvalues, an invariant subspace is spanned by some eigenvectors. The matrices A and B are unitarily irreducible since neither V^*V nor U^*U has a block of zeros in a position that allows orthogonal complementary eigenspaces.

By computation, it can be verified that the singular values of A and B are different. The fact that the singular values are different implies that A is unitarily similar to neither B nor B^T, and the proof is complete.

Alternatively, we can complete the proof without relying on a computation of the singular values as follows. The matrix A cannot be unitarily similar to B. If A and B were unitarily similar, Lemma 3.1 would imply that there was a diagonal matrix D_{uv} such that $V^*V = D_{uv}^*(U^*U)D_{uv}$. Such a D_{uv} cannot exist, since V^*V has zero in a position where U^*U has a non-zero.

Nor can A be unitarily similar to B^T. If it were, Lemma 3.1 would imply that there was a diagonal matrix D_{uv} such that $(V^*V)^{-T} = D_{uv}^*(U^*U)D_{uv}$. Such a D_{uv} cannot exist, since while it is easily verified that $(V^*V)^{-T}$ has a non-zero in every position, U^*U has zeros in some positions.

$\square$

Now we explain the idea that led to this example. First, choose a U^*U which is not an *arrowhead* matrix (all non-zeros are either on the diagonal or the last column or the last row) such that $U^*U \circ (U^*U)^{-T}$ is an arrowhead matrix. Then choose a V^*V that is an arrowhead such that $V^*V \circ (V^*V)^{-T} = U^*U \circ (U^*U)^{-T}$. This step involves solving a quadratic equation, hence the square roots in V^*V. The rest of the construction is trivial.

It is easy to find examples that satisfy some, but not all, of the conditions of Theorem 3.2. For example, for a suitable choice of C and D, the matrices $\mathrm{diag}(C, D)$ and $\mathrm{diag}(C, D^T)$ satisfy the first and last conditions.

Matrices of the kind described in Theorem 3.2 do not exist for $N = 1, 2, 3$. They exist, however, for all $N > 3$. The construction of examples for $N > 4$ is quite similar to the case $N = 4$. We omit the proof as it is clumsy.

We urge the reader to enter the matrices A and B in Theorem 3.2 into MATLAB and examine the singular values and the Frobenius norms of various functions of A and B.

References

[1] Anne Greenbaum, Norms of Functions of Matrices, Tech. Report 645, Courant Institute, NYU.

[2] Anne Greenbaum and L.N. Trefethen, Do the pseudospectra of a matrix determine its behaviour?, *Lin. Alg. Applics.*, to appear.

[3] R.A. Horn and C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).

[4] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, 1989).

[5] L.N. Trefethen, *Pseudospectra of Matrices*, in: J.C. Mason and M.G. Cox, eds., Algorithms for Approximation II, Chapman and Hall, London, 1990.