Acta Sci. Math. 44 (1982), 345—357

Doubly stochastic, unitary, unimodular, and complex orthogonal power embeddings

ROBERT C. THOMPSON and CHI-CHIN TUNG KUO

1. Introduction

We shall say that a finite matrix A is embedded in a larger finite matrix M if A is the leading principal submatrix of M, and we write $M \supset A$ or $A \subset M$. If $M^i \supset A^i$ for $i=1, 2, \ldots, k$, we say A is power embedded in M to exponent k. We also say that M is a dilation of A. Many years ago, in connection with unitary dilation theory for Hilbert space operators, E. EGERVÁRY [2] studied power embeddings of a contraction A into a unitary M. The objective of this note is to sharpen Egerváry's result and also to obtain analogous power embedding theorems into a doubly stochastic matrix, or into an integral unimodular matrix, or into a complex orthogonal matrix. The fact that more or less analogous theorems are obtainable suggests that various other parts of the presently existing rather extensive unitary dilation theory for infinite dimensional operators is capable of expansion in various directions. See, for example [1] and [5].

In each of our cases, the dilation M will turn out to exist if and only if it has at least $k\delta$ more rows then A, where δ is a measure of how far A is itself from the doubly stochastic, unimodular, unitary, or orthogonal state.

2. Doubly stochastic power embeddings

Let A be an $\alpha \times \alpha$ matrix with nonnegative entries. We consider whether it is possible to find a power embedding of A to exponent k into a doubly stochastic matrix M. (Thus M is to have real nonnegative entries with row and column sums equal to one.) Here k is fixed and specified in advance, and we wish also to know the size of the smallest dilation M.

Of course, a doubly stochastic dilation of A could exist only if A is a contraction in the sense of nonnegative matrices, that is, has each row and column sum at most one. In this case we say that A is a double stochastic contraction. So assume that A is

Received July 27, 1981.

a double stochastic contraction, set $A=[a_{ij}]_{1\leqq i,j\leqq \alpha}$, and let

$$d=\alpha-\sum_{i,j=1}^{\alpha} a_{ij}$$

measure the doubly stochastic deficiency of A. (Plainly $d\geqq 0$ with equality if and only if A is doubly stochastic.) Now take δ to be the least integer satisfying $\delta\geqq d$. This quantity δ is an integral measure of the doubly stochastic deficiency of A, with $\delta=0$ precisely when contraction A is doubly stochastic.

Theorem 1. *Let the nonnegative matrix A be a doubly stochastic contraction, and define the integral doubly stochastic deficiency δ of A as above. Then A possesses a doubly stochastic power embedding M to exponent k, i.e.,*

$$M^i \supset A^i \quad \textit{for} \quad 1\leqq i\leqq k, \tag{1}$$

with M having μ more rows than A, if and only if $\mu\geqq k\delta$.

The proof requires several lemmas.

Lemma 1. *Suppose that A is not doubly stochastic, that M is, and that* (1) *holds. Then, after a permutation similarity preserving A, M takes the form $M=[M_{ij}]_{0\leqq i,j\leqq k}$ with*

(a) $M_{00}=A$, *and the other diagonal blocks M_{ii} square,*
(b) $M_{ij}=0$ *whenever* $j\geqq i+2$,
(c) $M_{i0}=0$ *for all* i, $1\leqq i<k$,
(d) *Each row of each $M_{i,i+1}$ is nonzero, for all i with $2\leqq i\leqq k$. (Read $M_{k,k+1}$ as M_{k0}.)*

Proof. The proof is by induction on k, the case $k=1$ being trivial. Suppose the result established for k, and now assume that $M^i\supset A^i$ for $i=k+1$ also. Perform a permutation similarity on M, permuting only rows (and columns) that pass through M_{11}. Note that $M_{12}\neq 0$, since $M_{12}=0$ forces M_{11} to be doubly stochastic, hence $M_{01}=0$, and therefore forces $M_{00}=A$ to be doubly stochastic, a contradiction. (When $k=1$, M_{12} is M_{10}.) We chose our permutation similarity so that the nonzero rows in M_{12} are the last rows, i.e.,

$$M_{12}=\begin{bmatrix} M'_{13} \\ M'_{23} \end{bmatrix}$$

with $M'_{13}=0$ and with each row in M'_{23} nonzero. (Conceivably, block M'_{23} is vacuous, and when $k=1$ subscript 3 is read as 0.) Partition and renumber the blocks in

M in accord with this pattern:

$$M_{00} = M'_{00}, \qquad M_{01} = [M'_{01}, M'_{02}], \qquad M_{0j} = M'_{0,j+1},$$

$$M_{10} = \begin{bmatrix} M'_{10} \\ M'_{20} \end{bmatrix}, \quad M_{11} = \begin{bmatrix} M'_{11}, M'_{12} \\ M'_{21}, M'_{22} \end{bmatrix}, \qquad M_{1j} = \begin{bmatrix} M'_{1,j+1} \\ M'_{2,j+1} \end{bmatrix},$$

$$M_{i0} = M'_{i+1,0}, \quad M_{i1} = [M'_{i+1,1}, M'_{i+1,2}], \quad M_{ij} = M'_{i+1,j+1},$$

with square blocks M'_{11}, M'_{22}. Then M partitions as $M=[M'_{ij}]_{0\leqq i,j\leqq k+1}$ with $M'_{00}=A$. At this moment, it is conceivable that block row 1 is absent. We show that $M'_{02}=0$. The leading block in M^{k+1} is

$$A^{k+1}+M'_{02}M'_{23}\dots M'_{k,k+1}M'_{k+1,0}.$$

By hypothesis this equals A^{k+1}. Since $M'_{23}, \dots, M'_{k+1,0}$ each has all rows non-zero, we deduce that $M'_{02}=0$. If the block row labeled $i=1$ were absent, so would be block column i, hence $M'_{00}=A$ would be doubly stochastic, a contradiction. This completes the induction step.

Lemma 2. *Suppose that doubly stochastic $M=[M_{ij}]$ has the block form described in Lemma* 1, *where M_{ij} is $n_i \times n_j$. Then $d \leqq n_i$ for $1 \leqq i \leqq k$.*

Proof. If Q is a matrix, σQ will denote the sum of the entries of Q. Because M is doubly stochastic and has leading block row $A, M_{01}, 0, 0, \dots, 0$, we get $\sigma M_{01}=d$. Fix p, $1 \leqq p \leqq k$. Then, by columns,

$$\sigma[M_{ij}]_{1\leqq i\leqq k, 1\leqq j\leqq p} = n_1+\dots+n_p-d,$$

hence

$$\sigma[M_{ij}]_{1\leqq i\leqq p, 1\leqq j\leqq p} \leqq n_1+\dots+n_p-d,$$

therefore,

$$\sigma[M_{ij}]_{1\leqq i\leqq p, 1\leqq j\leqq p+1}-\sigma M_{p,p+1} \leqq n_1+\dots+n_p-d.$$

But

$$\sigma[M_{ij}]_{1\leqq i\leqq p, 1\leqq j\leqq p+1} = \sigma[M_{ij}]_{1\leqq i\leqq p, 0\leqq j\leqq k} = n_1+\dots+n_p.$$

Therefore $d \leqq \sigma M_{p,p+1} \leqq \sigma M_{p0}+\sigma M_{p1}+\dots+\sigma M_{p,k}=n_p$. Hence $d \leqq n_p$ as desired; $1 \leqq p \leqq k$. (Where necessary, read subscripts modulo $k+1$.)

Lemma 3. *If A possesses a doubly stochastic power embedding M, to exponent k, with M having μ more rows than A, then $\mu \geqq k\delta$.*

Proof. $\mu=n_1+\dots+n_k$ and $d \leqq n_i$ for each i. Since n_i is integral, $\delta \leqq n_i$ for all i, therefore $\mu \geqq k\delta$.

Lemma 4. *Given $\alpha\times\alpha$ contraction A, there exists an $(\alpha+\delta)\times(\alpha+\delta)$ doubly stochastic matrix*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Here δ is as before, $\delta\geqq d$.

Proof. We apply Ky Fan's criterion [4, see also 3] for the solvability of a mixed system of linear equalities and inequalities. Let the row sums of A be r_i, and the column sums be c_j; $1\leqq i, j\leqq\alpha$. The condition to be satisfied is that the matrix

$$\begin{bmatrix} 0 & B \\ C & D \end{bmatrix} \tag{2}$$

have row sums $1-r_i$ in the top block row, and 1 in the other block row, plus a corresponding column statement. These are the equalities to be considered. The inequalities are that the entries of B, C, D are to be nonnegative. We treat these entries as unknowns. If (p, q) is a position in B, C, or D, we introduce real dummy variables u_p, v_q and a real nonnegative dummy variable w_{pq}. Form a column vector u from the u_p, and a row vector v from the v_q. Form also a column vector $c=[1-r_1, 1-r_2, \ldots, 1-r_\alpha, 1, 1, \ldots, 1]^T$ displaying the proposed row sums in (2), and a row vector $r=[1-c_1, 1-c_2, \ldots, 1-c_\alpha, 1, 1, \ldots, 1]$ displaying the proposed column sums. Ky Fan's test for the solvability of our mixed system of equalities and inequalities amounts to this: We must show that the conditions

$$u_p+v_q+w_{pq}=0 \tag{3}$$

(u_p, v_q real, $w_{pq}\geqq 0$) for all (p, q) belonging to blocks B, C, D imply

$$(u, c)+(r, v)\leqq 0, \tag{4}$$

where $(\cdot\,,\,\cdot)$ is the standard inner product. This is easily done. Let U_0 be the maximum entry among $u_1, \ldots, u_\alpha$ and U_1 the maximum entry among $u_{\alpha+1}, \ldots, u_{\alpha+\delta}$. Similarly let V_0 be the maximum entry among $v_1, \ldots, v_\alpha$, and V_1 the maximum entry among $v_{\alpha+1}, \ldots, v_{\alpha+\delta}$. Noting that $\sum_1^\alpha(1-r_i)=\sum_1^\alpha(1-c_j)=d$, we get

$$(u, c)+(r, v)\leqq U_0d+U_1\delta+V_0d+V_1\delta =$$

$$=(U_0+V_1)d+(U_1+V_0)d+(U_1+V_1)(\delta-d)\leqq 0,$$

owing to (3) and $d\leqq\delta$.

Proof of Theorem 1. If A is doubly stochastic the necessity of $\mu\geqq k\delta$ is trivial since $\delta=0$. If A is not doubly stochastic, $\mu\geqq k\delta$ follows from Lemma 3. Conversely, let μ be any integer satisfying $\mu\geqq k\delta$. Construct the blocks B, C, D

described by Lemma 4. Now take M to be the direct sum of

$$\begin{bmatrix} A & B & 0 & 0 & \dots & 0 \\ 0 & 0 & I & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & 0 & \dots & I \\ C & D & 0 & 0 & \dots & 0 \end{bmatrix}$$

and a $(\mu-k\delta)$-square identity matrix. The identity matrices in the above block are $\delta\times\delta$. Then M is doubly stochastic and $M^i\supset A^i$ for $1\leqq i\leqq k$.

3. Unitary power embeddings

Our objective in this section is to sharpen Egerváry's theorem on unitary power embeddings. Let matrix A have complex elements. To be embeddible at all in a unitary matrix, each singular value of A must be $\leqq 1$, i.e., A must be a contraction. So assume that A is a contraction. The unitary deficiency δ of contraction A is now defined as the number (with multiplicity) of singular values of A strictly less than one. Thus $\delta=0$ if and only if A is unitary, and, in general for contraction A, δ is the rank of $I-AA^*$, also the rank of $I-A^*A$.

Theorem 2. *Let complex matrix A be a contraction, and define the unitary deficiency δ of A as above. Then A possesses a unitary power embedding M to exponent k, with M having μ more rows than A, if and only if $\mu\geqq k\delta$.*

When $k=1$, this Theorem is an easy special case of a known result [6] on singular values.

Lemma 5. *Suppose that A is not unitary, that M is, and that* (1) *holds with $k\geqq 2$. Then, after a unitary similarity preserving A, the matrix M takes the form $M=[M_{ij}]_{0\leqq i,j\leqq k}$ with*

(a) *$M_{00}=A$ and the other diagonal blocks square;*

(b) *each block is zero, except perhaps for M_{00}, M_{01}, M_{11}, M_{k0}, M_{k1}, and $M_{i,i+1}$ for $1\leqq i<k$;*

(c) *blocks $M_{23}, M_{34}, \dots, M_{k-1,k}$ are each unitary and $\delta\times\delta$;*

(d) *block M_{12} has δ columns and at least δ rows; block M_{k0} has δ rows, all linearly independent.*

Proof. We begin with $M=\begin{bmatrix} A & B \\ C & D\end{bmatrix}$. After a block diagonal similarity by a unitary matrix of the form diag (I, W), we preserve A and change C to WC. Choosing the last rows of W to be an orthonormal basis for the row space of C^*, we convert C to a matrix in which the first rows are zero and the last are linearly independent. So repartition M as

$$M = \begin{bmatrix} M_{00} & M_{01} & M_{02} \\ 0 & M_{11} & M_{12} \\ M_{20} & M_{21} & M_{22} \end{bmatrix}$$

with square diagonal blocks, $M_{00}=A$, and linearly independent rows in M_{20}. From $M^*M=I$ we get $M_{20}^*M_{20}=I-A^*A$. Since $I-A^*A$ has rank δ, this also is the rank of $M_{20}^*M_{20}$ and therefore of M_{20}. Hence M_{20} has δ rows. The leading block in M^2 is $A^2+M_{02}M_{20}$. Hence $M_{02}M_{20}=0$, and as M_{20} has independent rows, $M_{02}=0$. Orthogonality of columns then forces $M_{22}=0$. Since M_{22} is $\delta\times\delta$ and since $M_{02}=0$, M_{12} must have independent columns. It follows that M_{12} has at least δ rows. This completes the proof for $k=2$.

Suppose the result established for k, and now assume $M^{k+1}\supset A^{k+1}$. After a block diagonal unitary similarity preserving A, we may make the last rows of M_{12} independent and the remaining rows zero. Since M_{12} has δ columns, necessarily independent, this lower block in M_{12} is $\delta\times\delta$ and unitary. Now repartition as in the proof of Lemma 1. Then M'_{23} is $\delta\times\delta$ and M'_{12} possibly is vacuous. We must show that $M'_{02}=0$, M'_{12} has δ columns and at least δ rows, $M'_{21}=0$, $M'_{22}=0$, $M'_{k+1,2}=0$.

For simplicity, drop primes. Since $M^{k+1}\supset A^{k+1}$, we have

$$M_{02}M_{23}\ldots M_{k,k+1}M_{k+1,0}=0.$$

Linear independence of rows in $M_{23}, \ldots, M_{k+1,0}$ forces $M_{02}=0$. Orthogonality of columns forces $M_{21}=0$, $M_{22}=0$, $M_{k+1,2}=0$. If M_{12} were absent, M_{00} would be forced to be a direct summand of M, hence unitary. Therefore M_{12} is present, and as M_{22} is $\delta\times\delta$, M_{12} has δ columns, necessarily independent. Therefore it has at least δ rows.

Proof of Theorem 2. Suppose that a power embedding of A into a unitary matrix M exists, to exponent k. We wish to show that the number μ of additional rows in M satisfies $\mu\geqq k\delta$. If A is already unitary this is evident. Suppose A to be not unitary. If $k=1$ we have

$$M = \begin{bmatrix} A & M_{01} \\ M_{10} & M_{11} \end{bmatrix}$$

where $M_{10}^*M_{10}=I-A^*A$ has rank δ; therefore M_{10} has at least δ rows, hence $\mu\geqq\delta$. Now suppose $k\geqq 2$. Then M may be put in the form described in Lemma 5, with blocks $M_{12}, M_{23}, \ldots, M_{k0}$ each having at least δ rows. Therefore $\mu\geqq k\delta$.

Turning to the converse, since both $I-AA^*$ and $I-A^*A$ have rank δ, nonsingular matrices X and Y exist such that

$$I-AA^* = X\begin{bmatrix} I_\delta & 0 \\ 0 & 0\end{bmatrix}X^*, \quad I-A^*A = Y^*\begin{bmatrix} 0 & 0 \\ 0 & I_\delta\end{bmatrix}Y.$$

Set

$$B = X\begin{bmatrix} I_\delta \\ 0 \end{bmatrix}, \quad C = -[0, I_\delta]Y, \quad D = [0, I_\delta]YA^*X^{*-1}\begin{bmatrix} I_\delta \\ 0 \end{bmatrix},$$

and form the matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}. \tag{5}$$

We claim that this matrix is unitary. Certainly $AA^*+BB^*=I_\alpha$. Next, note that

$$X^{-1}AY^*\begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} = X^{-1}A(I-A^*A)Y^{-1} = X^{-1}(I-AA^*)AY^{-1} = \tag{6}$$
$$= \begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix} X^*AY^{-1}.$$

The definition of D shows that

$$\begin{bmatrix} D^* \\ 0 \end{bmatrix} = X^{-1}AY^*\begin{bmatrix} 0 \\ I_\delta \end{bmatrix}, \quad \text{and hence} \quad X\begin{bmatrix} I_\delta \\ 0 \end{bmatrix}D^* = AY^*\begin{bmatrix} 0 \\ I_\delta \end{bmatrix}.$$

This equation is the same as $AC^*+BD^*=0$.

Finally, we show that $CC^*+DD^*=I_\delta$. We have, using (6) at one point,

$$\begin{bmatrix} 0 & 0 \\ 0 & CC^* \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} Y\cdot Y^*\begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} = Y^{*-1}(I-A^*A)\cdot(I-A^*A)Y^{-1} =$$
$$= Y^{*-1}(I-2A^*A+A^*(AA^*)A)Y^{-1} =$$
$$= Y^{*-1}\left(I-2A^*A+A^*\left\{I-X\begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix}X^*\right\}A\right)Y^{-1} =$$
$$= Y^{*-1}\left(I-A^*A-A^*X\begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix}X^*A\right)Y^{-1} =$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} - Y^{*-1}A^*X\begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix}\cdot\begin{bmatrix} I_\delta & 0 \\ 0 & 0 \end{bmatrix}X^*AY^{-1} =$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix}YA^*X^{-1*}\cdot X^{-1}AY^*\begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} =$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ D & 0 \end{bmatrix}\begin{bmatrix} 0 & D^* \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & DD^* \end{bmatrix}.$$

Hence $CC^*+DD^*=I_\delta$, as claimed.

Therefore matrix (5) is unitary, and has δ more rows than A. This settles the case $k=1$. For $k>1$, we use these blocks B, C, D together with the construction in the proof of Theorem 1. This completes the proof of Theorem 2.

Let us note two additional facts: (i) If contraction A has real entries, this proof produces a real dilation M; (ii) If contraction A has quaternion entries, this proof produces a symplectic dilation M, i.e., a unitary M with quaternion entries.

It should also be noted that essentially this proof is already in the literaure; see [1].

4. Unimodular power embeddings

Now let $\alpha\times\alpha$ matrix A have entries from a commutative principal ideal domain R. We wish to embed A to exponent k into a unimodular matrix M. We take the unimodular deficiency δ of A to be the number of non-unit invariant factors of A. Then $\delta=0$ if and only if A is unimodular. We adopt the convention that any matrix A over R now is to be viewed as a contraction. (This is not quite as unnatural as it seems: if we were to permit matrices with entries from the field of fractions of R, by p-adic theory the contractions would be just those with entries in R.)

Theorem 3. *Let matrix A have entries in the principal ideal domain R. Then A possesses a unimodular power embedding M to exponent k, with M having μ more rows than A, if and only if $\mu\geqq k\delta$, where δ is the unimodular deficiency of A defined above.*

When $k=1$, this theorem is a special case of a known result [7] on invariant factors.

Lemma 6. *Assume that A is not unimodular, and that A is power embedded in unimodular M to exponent k. Then, after a unimodular similarity preserving A, M takes the form $M=[M_{ij}]_{0\leqq i,j\leqq k}$, where*

(a) $M_{00}=A$ *and each diagonal block M_{ii} is square,*

(b) $M_{ij}=0$ *if* $j\geqq i+2$,

(c) $M_{i0}=0$ *for* $1\leqq i<k$,

(d) *Each block $M_{12}, M_{23}, \ldots, M_{k0}$ has at least δ rows, with $M_{23}, \ldots, M_{k0}$ each having all rows independent and M_{12} at least δ independent rows.*

Proof. By induction on k. For $k=1$ we need only show that M_{10} has at least δ independent rows. After a block diagonal unimodular similarity of M preserving A, no generality is lost if M_{10} is cast into Hermite form. If it has fewer than δ nonzero rows, by a column Laplace expansion $\det M$ is a linear combination of $\alpha\times\alpha$ minors formed from the first α columns of M, each minor using at least $\alpha-\delta+1$ rows from M_{00}. Thus $\det M$ is a linear combination of $(\alpha-\delta+1)$-square minors from M_{00}. If s is the first nonunit invariant factor of $A=M_{00}$, each of these minors is divisible by s, so that s is a factor of $\det M$. This is impossible.

Now assume the result for k. By a block diagonal unimodular similarity of M

preserving A, we may cast M_{12} into row Hermite form, with the nonzero (and independent) rows last. Now repartition M as before. Then $M^{k+1}\supset A^{k+1}$ implies $M_{02}M_{23}\ldots M_{k,k+1}M_{k+1,0}=0$ and hence (by independence of rows) $M_{02}=0$. Let the blocks be $n_i\times n_j$. We must show that $\delta\leqq n_1$ and that M_{12} has at least δ independent rows. To see that $\delta\leqq n_1$, expand $\det M$ down its first $n_0=\alpha$ columns, with complementary minors coming from the last $n_1+\ldots+n_{k+1}$ columns. If a minor uses x rows from M_{00}, and if this minor is not to be divisible by s we must have $x\leqq n_0-\delta$. Also, it must use n_0-x rows from $M_{k+1,0}$.

Expand the complementary minor down the columns running through M_{01}. A nonzero minor in this expansion must use all the rows in M_{01} that were not used in M_{00}, and perhaps some rows from $M_{11}, \ldots, M_{k+1,1}$. The complement of this minor uses all the columns and some of the rows of

$$\begin{bmatrix} M_{12} & 0 & 0 & \ldots & 0 \\ M_{22} & M_{23} & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ M_{k2} & M_{k3} & \cdot & \ldots & M_{k,k+1} \\ M_{k+1,2} & M_{k+1,3} & \cdot & \ldots & M_{k+1,k+1} \end{bmatrix}.$$

There are $n_1+\ldots+n_{k+1}-(n_0-x)$ rows to select from to produce a nonzero minor, and $n_2+\ldots+n_{k+1}$ must be used. Consequently if there is to be a term in the expansion of $\det M$ not divisible by s, we must have $n_2+\ldots+n_{k+1}\leqq n_1+\ldots+n_{k+1}-(n_0-x)$. Thus $\delta\leqq n_0-x\leqq n_1$. If there were not δ independent rows in M_{12}, we could cast M_{12} into row Hermite form, and repeat the last argument with a smaller matrix M_{12} having less than δ rows.

Proof of Theorem 3. We first show that $\mu\geqq k\delta$. If A is unimodular this is clear. If not, M partitions into $n_i\times n_j$ blocks, $0\leqq i,j\leqq k$, with $\delta\leqq n_1, \delta\leqq n_2, \ldots, \delta\leqq n_k$. Hence $\mu\geqq k\delta$.

Conversely, we first produce a unimodular matrix of size $(\alpha+\delta)\times(\alpha+\delta)$:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Let $A=U\operatorname{diag}(I, S)V$, where U and V are unimodular, and S is diagonal with the nonunit invariant factors of A as diagonal elements. (The Smith form of A.) Here S is $\delta\times\delta$. Then let

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & I_\delta \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & S & I_\delta \\ 0 & I_\delta & 0 \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & I_\delta \end{bmatrix}.$$

This matrix is plainly unimodular and has δ more rows and columns than A. This settles the case $k=1$, and $k>1$ is now treated as in the proof of Theorem 1.

5. Complex orthogonal power embeddings

Let $\alpha\times\alpha$ matrix A now have complex entries. The closeness of A to (complex) orthogonality will be measured by the rank of δ of $I-AA^{\mathrm{T}}$. This also is the rank of $I-A^{\mathrm{T}}A$, because: the rank of $I-AA^{\mathrm{T}}$ is α minus the number of elementary divisors of AA^{T} belonging to eigenvalue 1. Since the elementary divisors of a matrix product AB belonging to a nonzero eigenvalue are also those of BA (Flander's Theorem), $A^{\mathrm{T}}A$ must have the same elementary divisors for 1 as AA^{T}. Of course, $\delta=0$ if and only if A is already orthogonal.

Theorem 4. *Let matrix A have complex entries. Then A possesses a complex orthogonal power embedding into M, to exponent k, with M having μ more rows than A, if and only if $\mu\geqq k\delta$, where δ is the orthogonal deficiency of A defined above.*

The proof of sufficiency is entirely analogous to the sufficiency proof in Theorem 2, changing * to $^{\mathrm{T}}$. Only the necessity needs proof. First we treat the case $k=1$. Let

$$M=\begin{bmatrix} A & M_{01} \\ M_{10} & M_{11} \end{bmatrix}.$$

Orthogonality demands that $M_{01}M_{01}^{\mathrm{T}}=I-AA^{\mathrm{T}}$, and hence M_{01} must have rank at least δ. Therefore it has at least δ columns.

The following lemma will be required below.

Lemma 7. *Let S be a $k\times n$ complex matrix with $SS^{\mathrm{T}}=I_k$. Then an $n\times n$ orthogonal matrix O exists with S as the last k rows.*

Proof. Plainly S has rank k, hence $k\leqq n$ and it has a $k\times k$ nonsingular submatrix. Let P be a permutation matrix such that SP has its initial $k\times k$ submatrix nonsingular. Set $SP=[S_1, S_2]$, with S_1 invertible. Now take

$$O=\begin{bmatrix} X & Y \\ S_1 & S_2 \end{bmatrix}P^{-1}$$

with $X=-YS_2^{\mathrm{T}}S_1^{-1\mathrm{T}}$. Then $XS_1^{\mathrm{T}}+YS_2^{\mathrm{T}}=0$, for any choice of Y. We require $XX^{\mathrm{T}}+YY^{\mathrm{T}}=I_{n-k}$ and this amounts to

$$Y[S_2^{\mathrm{T}}S_1^{-1\mathrm{T}}S_1^{-1}S_2+I_{n-k}]Y^{\mathrm{T}}=I_{n-k}. \tag{7}$$

Now $I_k=S_1S_1^{\mathrm{T}}+S_2S_2^{\mathrm{T}}=S_1[I_k+(S_1^{-1}S_2)(S_1^{-1}S_2)^{\mathrm{T}}]S_1^{\mathrm{T}}$. Hence -1 is not an eigenvalue of $(S_1^{-1}S_2)(S_1^{-1}S_2)^{\mathrm{T}}$, and therefore not of $(S_1^{-1}S_2)^{\mathrm{T}}(S_1^{-1}S_2)$. Thus $S_2^{\mathrm{T}}S_1^{-1\mathrm{T}}S_1^{-1}S_2+I_{n-k}$ is nonsingular, and hence (7) can be satisfied by some choice of Y. (Note that Y is square.) For this choice of Y, O is orthogonal, and the lemma is proved.

Now we handle the case $k=2$. Rename the blocks in M as

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Then $C^{\mathrm{T}}C = I - A^{\mathrm{T}}A = Y^{\mathrm{T}} \operatorname{diag}(0, I_\delta) Y$ for some invertible Y. Hence

$$(CY^{-1})^{\mathrm{T}}(CY^{-1}) = \operatorname{diag}(0, I_\delta).$$

Let S be the $\mu \times \delta$ matrix comprising the last δ columns of CY^{-1}. Then $S^{\mathrm{T}}S = I_\delta$, and by the lemma an orthogonal O exists of the form

$$O = \begin{bmatrix} R^{\mathrm{T}} \\ S^{\mathrm{T}} \end{bmatrix}.$$

Then $OCY^{-1} = \begin{bmatrix} Z & 0 \\ 0 & I_\delta \end{bmatrix}$, for some Z, and so $OC = \begin{bmatrix} ZY_1 \\ Y_2 \end{bmatrix}$, where Y_1, Y_2 arise from a partitioning of Y as $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$. Then

$$C^{\mathrm{T}}C = Y_1^{\mathrm{T}} Z^{\mathrm{T}} Z Y_1 + Y_2^{\mathrm{T}} Y_2 = (ZY_1)^{\mathrm{T}}(ZY_1) + Y^{\mathrm{T}} \begin{bmatrix} 0 & 0 \\ 0 & I_\delta \end{bmatrix} Y = (ZY_1)^{\mathrm{T}}(ZY_1) + C^{\mathrm{T}}C.$$

Therefore $(ZY_1)^{\mathrm{T}}(ZY_1) = 0$ and $Y_2^{\mathrm{T}} Y_2 = C^{\mathrm{T}}C$. Moreover Y_2 is $\delta \times \alpha$ with independent rows. Perform an orthogonal similarity on M by $\operatorname{diag}(I_\alpha, O)$. After this similarity, we partition M as

$$M = \begin{bmatrix} A & M_{01} & M_{02} \\ M_{10} & M_{11} & M_{12} \\ M_{20} & M_{21} & M_{22} \end{bmatrix}$$

with square diagonal blocks, $M_{20} = Y_2$ is $\delta \times \alpha$ with independent rows, and $M_{10} = ZY_1$, so that $M_{10}^{\mathrm{T}} M_{10} = 0$. If we pass to PMP^{T} with a block diagonal permutation matrix, we perform a permutation similarity on A and arrange that the initial $\delta \times \delta$ block in M_{20} is nonsingular, and still have $M_{10}^{\mathrm{T}} M_{10} = 0$.

We proceed to simplify the form of M. Let T be an as yet unspecified $\alpha \times \delta$ matrix. Observe that $M_{10}TT^{\mathrm{T}}M_{10}^{\mathrm{T}}$ is nilpotent (its square is zero), and therefore $I + M_{10}TT^{\mathrm{T}}M_{10}^{\mathrm{T}}$ is invertible. Choose a nonsingular X such that

$$X(I + M_{10}TT^{\mathrm{T}}M_{10}^{\mathrm{T}})X^{\mathrm{T}} = I,$$

then set $Y = -XM_{10}T$. Both X and Y depend on T. Now set

$$O = \begin{bmatrix} X & Y \\ T^{\mathrm{T}}M_{10}^{\mathrm{T}} & I_\delta \end{bmatrix}. \tag{8}$$

This matrix is orthogonal. Applying to M a block diagonal orthogonal similarity by diag (I_α, O), we get in the lower left positions

$$O\begin{bmatrix} M_{10} \\ M_{20} \end{bmatrix} = \begin{bmatrix} XM_{10}(I_\alpha - TM_{20}) \\ M_{20} \end{bmatrix}.$$

We now choose T, therefore also X and Y. Since the leading $\delta\times\delta$ submatrix in M_{20} is nonsingular, we may choose T so that

$$TM_{20} = \begin{bmatrix} I_\delta & \cdot \\ 0 & 0 \end{bmatrix}.$$

For this choice of T, the upper block in (8) has the form $[0, \cdot]$, where the 0 has δ columns. That is, after an orthogonal similarity preserving the structure of M, we may take M in the form

$$M = \begin{bmatrix} A & M_{01} & M_{02} \\ [0, M''_{10}] & M_{11} & M_{12} \\ [M'_{20}, M''_{20}] & M_{21} & M_{22} \end{bmatrix}$$

where M'_{20} is $\delta\times\delta$ and invertible. (The whole purpose of this reduction was to get a nonsingular block in M_{20} beneath a zero block in M_{10}.)

Now invoke the condition $M^2 \supset A^2$. This yields

$$M_{01}[0, M''_{10}] + M_{02}[M'_{20}, M''_{20}] = 0.$$

Because M'_{20} is invertible, we get $M_{02}=0$. And now, because M is orthogonal,

$$[0, M''_{10}]^{\mathrm{T}} M_{12} + [M'_{20}, M''_{20}]^{\mathrm{T}} M_{22} = 0,$$

yielding $M'^{\mathrm{T}}_{20} M_{22}=0$, whence $M_{22}=0$. Also, we now have $AA^{\mathrm{T}} + M_{01}M_{01}^{\mathrm{T}} = I_\alpha$, whence $M_{01}M_{01}^{\mathrm{T}}$ has rank δ, and thus M_{01} has at least δ columns. Hence M has at least 2δ more rows than A.

We have $M_{12}^{\mathrm{T}} M_{12} = I_\delta$, hence an orthogonal O exists with

$$O = \begin{bmatrix} Z \\ M_{12}^T \end{bmatrix},$$

for some Z. Then $OM_{12} = \begin{bmatrix} 0 \\ I_\delta \end{bmatrix}$, and a block diagonal orthogonal similarity of M by diag (I, O, I) preserves the block structure and converts M_{12} to $\begin{bmatrix} 0 \\ I_\delta \end{bmatrix}$. Repartitioning we now get

$$M = \begin{bmatrix} A & M_{01} & M_{02} & 0 \\ [0, M''_{10}] & M_{11} & M_{12} & 0 \\ [0, M''_{20}] & M_{21} & M_{22} & I_\delta \\ [M'_{30}, M''_{30}] & M_{31} & M_{32} & 0 \end{bmatrix}$$

with M'_{30} $\delta\times\delta$ and nonsingular. Orthogonality implies M''_{20}, M_{21}, M_{22} are all zero.

We now continue by induction, analogous to the proof of Theorem 2. For example, $M^3 \supset A^3$ now implies $M_{02}=0$ (using $M^2 \supset A^2$), whence $M_{32}=0$, and a splitting of M_{12} can be obtained, etc. This completes the proof of Theorem 4.

Comment. Each theorem above is of the following type: Given a semigroup G of matrices of specified size $n \times n$, and a fixed matrix A, how large must n be so that M exists in G with $M^i \supset A^i$ for $i=1, \ldots, k$. The same question can be formulated for other semigroups. For example, if G is the full linear group, then M must have at least $k\delta$ more rows than A, where δ is the nullity of A. We omit the proof.

References

[1] J. Bognár, *Indefinite inner product spaces*, Ergebnisse der Mathematik, Band 78, Springer-Verlag (Berlin—Heidelberg—New York, 1974).

[2] E. Egerváry, On the contractive linear transformations of n-dimensional vector space, *Acta Sci. Math.*, **15** (1954), 178—182.

[3] Ky Fan, On systems of linear inequalities, in: *Linear inequalities and related systems* (H. Kuhn and A. W. Tucker editors), Annals of Math. Studies, no. 38 (Princeton, 1956); 99—156.

[4] Ky Fan, Two applications of a consistency condition for systems of linear inequalities, *Linear Algebra and Applications*, **11** (1975), 171—180.

[5] B. Sz.-Nagy and C. Foiaş, *Harmonic analysis of operators on Hilbert space*, North Holland (Amsterdam, 1970).

[6] R. C. Thompson, Principal submatrices IX: Introducing inequalities for singular values, *Linear Algebra and Applications*, **5** (1972), 1—12.

[7] R. C. Thompson, Interlacing inequalities for invariant factors, *Linear Algebra and Applications*, **24** (1970), 1—31.

ALGEBRA INSTITUTE, MATHEMATICS DEPARTMENT
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CALIFORNIA 93106, USA