IMBEDDING CONDITIONS FOR HERMITIAN AND NORMAL MATRICES

KY FAN AND GORDON PALL

1. Imbedding theorems. Let A, B be two square matrices with complex coefficients, of respective orders n and m, where $n \geqslant m$. We shall say that B is *imbeddable* in A if there exists a unitary matrix U of order n such that U^*AU contains B as a principal submatrix. In other words, B is said to be imbeddable in A if there exists a matrix V of type $n \times m$ such that $V^*V = I_m$ (= the identity matrix of order m) and $V^*AV = B$.

For Hermitian matrices, the following result holds:

THEOREM 1. *Let A, B be two Hermitian matrices of respective orders n and m, where $n \geqslant m$. Let $\alpha_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_n$ and $\beta_1 \geqslant \beta_2 \geqslant \ldots \geqslant \beta_m$ be the characteristic roots of A and B respectively. Then a necessary and sufficient condition for B to be imbeddable in A is that inequalities*

$$(1) \qquad \alpha_i \geqslant \beta_i, \ \alpha_{n-i+1} \leqslant \beta_{m-i+1} \qquad (1 \leqslant i \leqslant m)$$

be fulfilled.

The necessity part of Theorem 1 is well known (**2**, p. 75). Inequalities (1) had already been given by Cauchy (**1**, p. 187) for real symmetric matrices. To the best of our knowledge, no proof of the sufficiency part has been published except for $m = 1$ or n. Dr. A. J. Hoffman, to whom we described our proof in the summer of 1954, kindly sent us an unpublished proof given by H. Wielandt in 1953 and based on quite different ideas.

For normal matrices, we have the following result:

THEOREM 2. *Let A, B be two normal matrices of respective orders n and $n - 1$. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ and $\beta_1, \beta_2, \ldots, \beta_{n-1}$ be the characteristic roots of A and B respectively. Renumber them so that $\alpha_1, \alpha_2, \ldots, \alpha_q$ are each distinct from $\beta_1, \beta_2, \ldots, \beta_{q-1}$, while $\alpha_j = \beta_{j-1}$ for $q + 1 \leqslant j \leqslant n$. Then a necessary and sufficient condition for B to be imbeddable in A is that the $2q - 1$ points $\alpha_1, \alpha_2, \ldots, \alpha_q, \beta_1, \beta_2, \ldots, \beta_{q-1}$ in the complex plane shall be distinct, collinear, and that every segment on this line limited by two adjacent α_i's $(1 \leqslant i \leqslant q)$ shall contain one $\beta_j (1 \leqslant j \leqslant q - 1)$.*

This result generalizes the case $m = n - 1$ of Theorem 1. It may appear surprising that a generalization to normal matrices is possible. For one thing, the roots are now complex, and since complex numbers have no natural ordering, inequalities (1) might not be extendible.

Received July 6, 1956. Work prepared in part under a National Bureau of Standards contract with the American University, sponsored by the Office of Scientific Research, USAF; and in part at the University of Notre Dame, under the sponsorship of the National Science Foundation.

We observe that the essential result in Theorem 1 is the case $m = n - 1$, in which case (1) becomes

$$\alpha_1 \geqslant \beta_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_{n-1} \geqslant \beta_{n-1} \geqslant \alpha_n. \tag{2}$$

The result for this case is easily extended to the case $m < n - 1$. In fact, if (1) holds, intermediary sequences of characteristic roots can be inserted such that two consecutive sequences are interlaced similarly to (2). Then, by the result for $m = n - 1$, there exists a chain of Hermitian matrices, with orders increasing by unity, such that each is imbeddable in the next.

For normal matrices, Theorem 2 deals only with the case $m = n - 1$. The case $m < n - 1$ seems to involve additional complications. If a normal matrix B of order $m < n - 1$ is imbeddable in a normal matrix A of order n, it is not true in general that there exists a chain of normal matrices, beginning with B and ending with A, with orders increasing by unity, such that each is imbeddable in the next. As an example, let A, B be diagonal matrices:

$$A = \operatorname{diag}\{0, 1, i, 1 + i\}, \quad B = \operatorname{diag}\left\{\frac{5 + 8i}{10}, \frac{5 + 2i}{10}\right\}.$$

If we take

$$V = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 & 2 \\ 1 & -2 \\ 2 & -1 \\ 2 & 1 \end{bmatrix},$$

then $V^*V = I_2$ and $V^*AV = B$; so B is imbeddable in A. By Theorem 2, any normal matrix C of order 3 imbeddable in A must have two of 0, 1, i, $1 + i$ as characteristic roots and its third root on the segment joining the remaining two. But by Theorem 2, B cannot be imbedded in any such matrix C.

Because of interest in extensions of Witt's theorem on quadratic forms in a field, the following corollary is worth noting: If a normal matrix with characteristic roots $\alpha_1, \beta_2, \ldots, \beta_{n-1}$ can be imbedded in one with roots $\alpha_1, \alpha_2, \ldots, \alpha_n$, then the same is true of the respective matrices with roots $\beta_2, \ldots, \beta_{n-1}$ and $\alpha_2, \ldots, \alpha_n$. Another interesting fact is as follows: A necessary and sufficient condition that *all* matrices imbeddable in a normal matrix A be normal is that the characteristic roots of A shall be collinear, or (what is the same thing) that A have the form $w(I + e^{i\theta}H)$ with H Hermitian, w a complex number and θ real. The sufficiency is evident:

$$V^*(I + e^{i\theta}H)V = I_m + e^{i\theta}K,$$

as $K = V^*HV$ is Hermitian. To prove the necessity notice that if N is normal then $w_1N + w_2I$ is normal (w_1 and w_2 complex numbers). Hence it suffices to consider $A = \operatorname{diag}\{0, 1, a + bi\}$, $b \neq 0$. But if

$$V^* = \begin{bmatrix} k & k & k \\ h & -h & 0 \end{bmatrix}, \quad h = 1/\sqrt{2}, \ k = 1/\sqrt{3},$$

the matrix

$$B = V^*AV = \begin{bmatrix} \frac{1}{3}(1+a+bi) & -hk \\ -hk & \frac{1}{2} \end{bmatrix}$$

is not normal.

2. Proof of the sufficiency part of Theorem 1. According to a remark made above it suffices to consider $m = n - 1$. Since the properties in question are invariant under unitary transformations we need only prove the following: If real numbers $\alpha_i (1 \leqslant i \leqslant n)$ and $\beta_j (1 \leqslant j \leqslant n-1)$ satisfy (2), then there exist $n-1$ complex numbers $z_i (1 \leqslant i \leqslant n-1)$ and a real number γ such that the Hermitian matrix

$$(3)\qquad \begin{bmatrix} \gamma & z_1 & z_2 & z_3 & \ldots z_{n-1} \\ \bar{z}_1 & \beta_1 & 0 & 0 & \ldots 0 \\ \bar{z}_2 & 0 & \beta_2 & 0 & \ldots 0 \\ \bar{z}_3 & 0 & 0 & \beta_3 & \ldots 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots\cdot \\ \bar{z}_{n-1} & 0 & 0 & 0 & \ldots \beta_{n-1} \end{bmatrix}$$

has $\{\alpha_i\}$ as characteristic roots. To show this, we may assume

$$(4)\qquad \alpha_1 > \beta_1 > \alpha_2 > \beta_2 > \ldots > \beta_{n-1} > \alpha_n$$

instead of (2). In fact, if $\beta_j = \alpha_j$ or α_{j+1}, we can choose $z_j = 0$ and then work with a matrix of order decreased by unity. Since the characteristic polynomial of the matrix (3) is

$$\left\{\lambda - \gamma - \sum_{j=1}^{n-1} \frac{|z_j|^2}{\lambda - \beta_j}\right\} \cdot \prod_{j=1}^{n-1} (\lambda - \beta_j),$$

and in view of the strict inequalities (4), the requirement that $\{\alpha_i\}$ shall be the characteristic roots of the matrix (3) is equivalent to

$$\gamma + \sum_{j=1}^{n-1} \frac{|z_j|^2}{\alpha_i - \beta_j} = \alpha_i \qquad (1 \leqslant i \leqslant n).$$

Hence it remains to prove the following

LEMMA. *Under the hypothesis* (4), *the system of n linear equations*

$$(5)\qquad \frac{x_1}{\alpha_i - \beta_1} + \frac{x_2}{\alpha_i - \beta_2} + \ldots + \frac{x_{n-1}}{\alpha_i - \beta_{n-1}} + x_n = \alpha_i \qquad (1 \leqslant i \leqslant n)$$

has a unique solution $x_1, x_2, \ldots, x_n$, *and this solution satisfies* $x_k > 0$ $(1 \leqslant k \leqslant n-1)$.

Proof. Let Δ denote the determinant of the coefficient matrix of the system (5). Let Δ_k denote the determinant obtained from Δ by replacing the kth column by the column of quantities α_i. Then we are to prove that $\Delta \neq 0$ and that Δ_k/Δ is positive for $1 \leqslant k \leqslant n-1$.

This will be seen from certain expressions identically equal to Δ or Δ_k. In deriving these expressions we shall regard the α's and β's as variables. First we write

$$\Delta = P/Q, \tag{6}$$

where

$$Q = \prod_{1\leqslant i\leqslant n}\ \prod_{1\leqslant j\leqslant n-1} (\alpha_i - \beta_j), \tag{7}$$

and where P is a polynomial in the $2n - 1$ variables α_i and β_j. Since two rows, or two columns, of Δ become equal when two α's, or two β's, are set equal, every difference $\alpha_i - \alpha_j$ $(1 \leqslant i < j \leqslant n)$ and $\beta_i - \beta_j$ $(1 \leqslant i < j \leqslant n - 1)$ is a factor of P. The number of these factors being

$$\binom{n}{2} + \binom{n-1}{2} = (n-1)^2,$$

P being homogeneous of degree $(n - 1)^2$, we must have

$$P = c \prod_{1\leqslant i<j\leqslant n} (\alpha_i - \alpha_j) \cdot \prod_{1\leqslant i<j\leqslant n-1} (\beta_i - \beta_j),$$

where c is a constant. To determine c we may choose

$$\beta_j = j\ (1 \leqslant j \leqslant n-1),\quad \alpha_i = i + \epsilon\ (1 \leqslant i \leqslant n-1),\quad \alpha_n = 0.$$

After this substitution, if we multiply by ϵ each of the first $n - 1$ columns of Δ, then from the complete expansion of the determinant it is easily seen that $\epsilon^{n-1}\Delta \to 1$ as $\epsilon \to 0$. In the polynomial P in ϵ, the constant term is

$$c(n-1)! \prod_{1\leqslant i<j\leqslant n-1} (i-j)^2.$$

Also, we observe that ϵ^{n-1} is a factor of the polynomial Q in ϵ, and that the constant term in the polynomial Q/ϵ^{n-1} is

$$(-1)^{\frac{1}{2}(n-1)n}(n-1)! \prod_{1\leqslant i<j\leqslant n-1} (i-j)^2.$$

So we have $\epsilon^{n-1}\ P/Q \to c(-1)^{\frac{1}{2}(n-1)n}$, as $\epsilon \to 0$. Hence $c = (-1)^{\frac{1}{2}(n-1)n}$ and

$$P = (-1)^{\frac{1}{2}(n-1)n} \prod_{1\leqslant i<j\leqslant n} (\alpha_i - \alpha_j) \cdot \prod_{1\leqslant i<j\leqslant n-1} (\beta_i - \beta_j). \tag{8}$$

Similarly, we write

$$\Delta_k = P_k/Q_k \qquad (1 \leqslant k \leqslant n-1). \tag{9}$$

where

$$Q_k = \prod_{1\leqslant i\leqslant n}\ \prod_{\substack{1\leqslant j\leqslant n-1\\ j\neq k}} (\alpha_i - \beta_j), \tag{10}$$

and

$$P_k = c_k \prod_{1\leqslant i<j\leqslant n} (\alpha_i - \alpha_j) \cdot \prod_{\substack{1\leqslant i<j\leqslant n-1\\ i,\, j\neq k}} (\beta_i - \beta_j).$$

To determine the constant c_k, we use the same substitution as above. If we

https://doi.org/10.4153/CJM-1957-036-1 Published online by Cambridge University Press

multiply by ϵ each column of Δ_k except the kth and nth, then we get $\epsilon^{n-2}\Delta_k \to k$, as $\epsilon \to 0$. One also verifies that

$$\epsilon^{n-2} P_k/Q_k \to k c_k (-1)^{k+\frac{1}{2}n(n-3)}, \qquad \epsilon \to 0.$$

So we have

$$(11) \qquad P_k = (-1)^{k+\frac{1}{2}n(n-3)} \prod_{1 \leqslant i < j \leqslant n} (\alpha_i - \alpha_j) \cdot \prod_{\substack{1 \leqslant i < j \leqslant n-1 \\ i, j \neq k}} (\beta_i - \beta_j).$$

From (6), (7), (8) and (4), it is clear that $\Delta \neq 0$. Collecting our results we obtain, after simplification,

$$(12) \qquad \Delta_k/\Delta = - \prod_{1 \leqslant i \leqslant n} (\alpha_i - \beta_k) \Big/ \prod_{\substack{1 \leqslant j \leqslant n-1 \\ j \neq k}} (\beta_j - \beta_k) \qquad (1 \leqslant k \leqslant n-1),$$

which is positive, by (4). This completes the proof.

From the above proof, it is clear that in the case of real symmetric matrices A, B, Theorem 1 remains valid if we require U to be orthogonal in our definition of imbedding.

3. Proof of Theorem 2. We first prove the necessity part. Let A, B be two normal matrices of orders n and $n-1$ respectively. Let α_i $(1 \leqslant i \leqslant n)$ be the characteristic roots of A, and let β_j $(1 \leqslant j \leqslant n-1)$ be those of B. Assume that B is imbeddable in A. Then there exists a unitary matrix U of order n such that U^*AU is of the form

$$(13) \qquad U^*AU = \begin{bmatrix} \gamma & z_1 & z_2 & z_3 & \dots z_{n-1} \\ \bar{w}_1 & \beta_1 & 0 & 0 & \dots 0 \\ \bar{w}_2 & 0 & \beta_2 & 0 & \dots 0 \\ \bar{w}_3 & 0 & 0 & \beta_3 & \dots 0 \\ \cdot & \cdot & \cdot & \cdot & \dots \cdot \\ \bar{w}_{n-1} & 0 & 0 & 0 & \dots \beta_{n-1} \end{bmatrix}.$$

The fact that U^*AU is normal can be expressed by

$$(14) \qquad |z_j| = |w_j| \qquad (1 \leqslant j \leqslant n-1),$$

$$(15) \qquad z_j \bar{z}_k = w_j \bar{w}_k \qquad (1 \leqslant j \leqslant n-1,\ 1 \leqslant k \leqslant n-1),$$

$$(16) \qquad (\beta_j - \gamma) w_j = \overline{(\beta_j - \gamma)} z_j \qquad (1 \leqslant j \leqslant n-1).$$

From (15), we have $(z_j \bar{w}_j)(z_k \bar{z}_k) = (z_k \bar{w}_k)(w_j \bar{w}_j)$, so among the $n-1$ numbers $z_j \bar{w}_j$ $(1 \leqslant j \leqslant n-1)$, the non-zero ones have the same amplitude. Designating this common amplitude by 2θ, we have by (14):

$$(17) \qquad z_j = e^{i2\theta} w_j \qquad (1 \leqslant j \leqslant n-1).$$

We may assume that $z_q, z_{q+1}, \dots, z_{n-1}$ are all those z_j's which are zero (if they exist). If $1 \leqslant j \leqslant q-1$, then by (16), either $\beta_j - \gamma = 0$, or

$$\operatorname{amp}(\beta_j - \gamma) \equiv \theta \pmod{\pi}.$$

In either case, we can set

$$\beta_j = \gamma + e^{i\theta} b_j \qquad (1 \leqslant j \leqslant q-1), \tag{18}$$

where b_j is real. Put

$$u_j = e^{-i\theta} z_j \qquad (1 \leqslant j \leqslant q-1). \tag{19}$$

Then by (17),

$$\bar{u}_j = e^{-i\theta} \bar{w}_j \qquad (1 \leqslant j \leqslant q-1). \tag{20}$$

Introduce the Hermitian matrix of order q

$$H = \begin{bmatrix} 0 & u_1 & u_2 & u_3 & \ldots u_{q-1} \\ \bar{u}_1 & b_1 & 0 & 0 & \ldots 0 \\ \bar{u}_2 & 0 & b_2 & 0 & \ldots 0 \\ \cdot & \cdot & \cdot & \cdot & \ldots\ldots \\ \bar{u}_{q-1} & 0 & 0 & 0 & \ldots b_{q-1} \end{bmatrix} \tag{21}$$

and let

$$C = \gamma I_q + e^{i\theta} H, \quad D = \operatorname{diag}\{\beta_q, \beta_{q+1}, \ldots, \beta_{n-1}\}. \tag{22}$$

Then according to (18), (19) and (20), we have

$$U^*AU = \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix}. \tag{23}$$

Let $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_q$ be the characteristic roots of the Hermitian matrix H. Then by (22) and (23), the characteristic roots $\{\alpha_i\}$ of A (also of U^*AU) can be so renumbered that

$$\begin{cases} \alpha_j = \gamma + e^{i\theta} a_j & (1 \leqslant j \leqslant q), \\ \alpha_j = \beta_{j-1} & (q+1 \leqslant j \leqslant n). \end{cases} \tag{24}$$

Now we renumber $\beta_1, \beta_2, \ldots, \beta_{q-1}$ in such a way that the corresponding $b_1, b_2, \ldots, b_{q-1}$ (see (18)) are arranged in decreasing order. Then, since $a_1, a_2, \ldots, a_q$ are the characteristic roots of H, we have the interlacing inequalities

$$a_1 \geqslant b_1 \geqslant a_2 \geqslant b_2 \geqslant \ldots \geqslant b_{q-1} \geqslant a_q. \tag{25}$$

Relations (18), (24) and (25) together express precisely the condition stated in the theorem.

To prove the sufficiency, let complex numbers α_i $(1 \leqslant i \leqslant n)$ and β_j $(1 \leqslant j \leqslant n-1)$ satisfy the condition stated in the theorem. Then for some complex number γ and real numbers

$$\theta;\ a_1, a_2, \ldots, a_q;\ b_1, b_2, \ldots, b_{q-1},$$

we have the relations (18), (24) and (25). From (25) and Theorem 1, there exists a Hermitian matrix H of order q, with characteristic roots $a_1, a_2, \ldots, a_q$, and containing $\operatorname{diag}\{b_1, b_2, \ldots, b_{q-1}\}$ as a principal submatrix. Using this H and the numbers γ, θ appearing in (18), (24), we form the normal matrix

$$C = \gamma I_q + e^{i\theta} H.$$

Let $D = \text{diag}\{\beta_q, \beta_{q+1}, \ldots, \beta_{n-1}\}$. Then

$$\begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix}$$

is a normal matrix of order n, and by (24), its characteristic roots are $\alpha_1, \alpha_2, \ldots, \alpha_n$. Since H contains $\text{diag}\{b_1, b_2, \ldots, b_{q-1}\}$ as a principal submatrix, so by (18),

$$\begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix}$$

contains $\text{diag}\{\beta_1, \beta_2, \ldots, \beta_{n-1}\}$ as a principal submatrix.

This shows that every normal matrix of order $n - 1$ with characteristic roots β_j $(1 \leqslant j \leqslant n - 1)$ is imbeddable in any normal matrix of order n with characteristic roots α_i $(1 \leqslant i \leqslant n)$.

References

1. A. Cauchy, *Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes;* Oeuvres complètes, Second Ser., IX, 174–195.
2. H. L. Hamburger and M. E. Grimshaw, *Linear Transformations in n-dimensional Vector Space* (Cambridge, 1951).

University of Notre Dame and American University

Illinois Institute of Technology