Can. J. Math., Vol. XXXIX, No. 6, 1987, pp. 1446-1458

CONGRUENCES AND NORMS OF HERMITIAN MATRICES

STEPHEN PIERCE AND LEIBA RODMAN

1. Introduction. Two complex hermitian $n \times n$ matrices A and B are *congruent* if $S^*AS = B$ for some invertible $n \times n$ matrix S (with complex entries). The matrix S is called a *congruence matrix*. Given congruent hermitian matrices A and B, a congruence matrix is, of course, not unique. For instance, if $A = B$ then one can take $S = \alpha I$ with $|\alpha| = 1$, as well as any other matrix satisfying $S^*AS = A$. However, here the choice $S = I$ seems naturally to be best possible in the sense that when applied to an n-dimensional column vector it produces no distortion or movement of the vector at all. We shall measure the distortion (or movement) of the vector $x \in \mathbf{C}^n$ under an $n \times n$ invertible matrix A in terms of $\|x - Ax\|$, where the norm is euclidean. Then the distortion produced by A is $\|I - A\|$, with the induced operator norm.

For given congruent hermitian matrices A and B define

$$\Omega(A, B) = \inf\{ \|I - S\| \,|S^*AS = B\}.$$

Thus $\Omega(A, B)$ measures the minimal distortion induced by congruence matrices which carry the congruence between A and B. In this paper we study the relationship between $\Omega(A, B)$ and the distance $\|A - B\|$. It turns out that, roughly speaking, $\Omega(A, B)$ grows in magnitude no faster than $\|A - B\|$. Similar results concerning unitary congruence are obtained as well.

In the last section we state a conjecture concerning simultaneous congruence of several hermitian matrices.

We state all results for the case of complex matrices only; similar results, with the same proofs, are valid also for matrices over the real field.

This investigation was inspired by analogous results concerning similarity of matrices (see [**2**], [**3**]).

2. Hermitian matrices congruent to a given matrix. We state and prove here the main results concerning the behavior of $\Omega(A, B)$.

Received April 3, 1986 and in revised form January 7, 1987. The work of the first author was partially supported by National Research Council of Canada Grant A-7862, and by National Science Foundation Grant DMS-8601959. The work of the second author was partially supported by National Science Foundation Grant DMS-8501794.

THEOREM 2.1. *Let A be an $n \times n$ hermitian matrix. Then there is a positive K such that*

$$\Omega(A, B) \leqq K\|B - A\|$$

for every B which is congruent to A.

The reader should note the global nature of the result (namely, matrices B which are congruent to, but far away from A, are compared to A along with the matrices which are congruent and close to A). In fact, the proof of Theorem 2.1 will be given by proving first the corresponding local version.

In connection with Theorem 2.1 we emphasize that in general $\Omega(A, B) \neq \Omega(B, A)$ (see Example 3.3 in Section 3); moreover, as we shall see in Example 3.1, Theorem 2.1 is false if $\Omega(A, B)$ is replaced by $\Omega(B, A)$.

It turns out that the constant K in Theorem 2.1 is uniformly bounded as long as the non-zero eigenvalues of A stay away from zero and do not grow indefinitely. To formalize this statement, for every non-zero hermitian matrix A define

$$\kappa(A) = \max\{ |\lambda|, |\lambda|^{-1}\},$$

where the maximum is taken over all non-zero eigenvalues λ of A. For a given positive number α define S_α to be the set of all $n \times n$ hermitian matrices A such that $\kappa(A) \leqq \alpha$. We have the following.

THEOREM 2.2. *For every $\alpha > 0$ there exists $K > 0$ such that*

$$\Omega(A, B) \leqq K\|B - A\|$$

for every $A \in S_\alpha$ and every B which is congruent to A.

For the proof of Theorem 2.1 it is convenient to establish several lemmas (which are particular cases of Theorem 2.1) first.

LEMMA 2.3. *Let A be positive definite Hermitian. Then there exist positive real numbers K and ϵ such that for every positive definite Hermitian matrix B with $\|B - A\| < \epsilon$ we have*

$$\max(\Omega(A, B), \Omega(B, A)) \leqq K\|B - A\|. \tag{2.1}$$

Proof. We first perform a reduction to verify (2.1). We will show that it suffices to assume that $A = I$. Suppose our lemma is verified for $A = I$, and now let A be an arbitrary positive definite Hermitian matrix.

We write $A = D^2$, where D is the positive definite Hermitian square root of A. If $\|B - D^2\| < \epsilon$, then

$$\|D^{-1}BD^{-1} - I\| < \epsilon\|D^{-1}\|^2.$$

Thus we may choose K and ϵ such that $\|B - D^2\| < \epsilon$ implies the existence of an X such that

https://doi.org/10.4153/CJM-1987-068-9 Published online by Cambridge University Press

(2.2) $X^*D^{-1}BD^{-1}X = I$ and $\|I - X\| \leqq K\|D^{-1}BD^{-1} - I\|$.

Now let X be as above and let $E = D^{-1}XD$ and note that $E^*BE = A$. Hence, if $\|B - D^2\| < \epsilon$, we have

$$\|I - E\| < \|D^{-1}\| \, \|D\| \, \|I - X\|$$
$$\leqq K\|D^{-1}\| \, \|D\| \, \|D^{-1}BD^{-1} - I\| \leqq K\|D^{-1}\|^3\|D\| \, \|B - A\|.$$

Thus $\|B - A\| < \epsilon$ implies the existence of an E such that

$$E^*BE = A \quad \text{and} \quad \|I - E\| < L\|B - A\|,$$

where $L = K\|D^{-1}\|^3\|D\|$. This verifies the reduction for $\Omega(A, B)$.

To check $\Omega(B, A)$, note that

$$(2.3) \quad \Omega(B, A) = \inf\{ \|I - Y\| \, | Y^*AY = B\}$$
$$\leqq \inf\{ \|Y\| \, \|I - Y^{-1}\| \, |A = Y^{*-1}BY^{-1}\}.$$

Now observe that by the already proved part of this reduction, we can find ϵ, $K > 0$ such that

$$\Omega(A, B) \leqq K\|B - A\|$$

for any matrix B congruent to A and satisfying $\|B - A\| < \epsilon$. So for such B there exists an invertible matrix Y with the properties that $A = Y^{*-1}BY^{-1}$ and

$$\|I - Y^{-1}\| \leqq K\|B - A\|.$$

Let $\epsilon' = \min(\epsilon, (2K)^{-1})$. Then, assuming in addition that $\|B - A\| < \epsilon'$ we easily see that $\|Y\| \leqq 2$. Compare with (2.3) to deduce that

$$\Omega(B, A) \leqq 2K\|B - A\|,$$

and the reduction to the case $A = I$ is completed.

From now on we assume that $A = I$. Further, by replacing B with U^*BU, where U is unitary, it is sufficient to prove that (2.1) holds for diagonal B's. Let

$$B = \operatorname{diag}(b_1, \dots, b_n), \ b_i > 0.$$

Then

$$\operatorname{diag}(\sqrt{b_1}, \dots, \sqrt{b_n}) \cdot A \cdot \operatorname{diag}(\sqrt{b_1}, \dots, \sqrt{b_n}) = B,$$

and hence

$$\Omega(I, B) \leqq \max(|1 - \sqrt{b_1}|, \dots, | 1 - \sqrt{b_n}|).$$

On the other hand,

$$\|B - I\| = \max(|1 - b_1|, \dots, | 1 - b_n|).$$

But

$$\frac{|1 - \sqrt{b_i}|}{|1 - b_i|} = \frac{1}{1 + \sqrt{b_i}} < 1,$$

and $\Omega(I, B) < \|B - I\|$ follows.

Further,

$$\Omega(B, I) \leqq \max\left(\left|1 - \frac{1}{\sqrt{b_1}}\right|, \ldots, \left|1 - \frac{1}{\sqrt{b_n}}\right|\right)$$

and

$$\frac{\left|1 - \dfrac{1}{\sqrt{b_i}}\right|}{|1 - b_i|} = \frac{1}{\sqrt{b_i}(1 + \sqrt{b_i})}$$

so, if b_i is sufficiently close to 1, (we take $\|B - I\| < \epsilon$)

$$\Omega(B, I) \leqq 2\|B - I\|.$$

LEMMA 2.4. *For every $n \times n$ hermitian matrix A there exist positive constants K and ϵ such that*

$$\max(\Omega(B, A), \Omega(A, B)) \leqq K\|B - A\|,$$

where B is any matrix which is congruent to A and satisfies $\|B - A\| < \epsilon$.

Proof. We shall prove only the inequality

$$\Omega(A, B) \leqq K\|B - A\|$$

(the second inequality

$$\Omega(B, A) \leqq K\|B - A\|$$

can be deduced from the first as in the reduction to the case $A = I$ in the proof of Lemma 2.3).

Without loss of generality we may assume that

$$(2.4) \qquad A = \begin{bmatrix} I_m & 0 & 0 \\ 0 & -I_p & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Let $\epsilon > 0$ be so small that for any hermitian matrix

$$(2.5) \qquad B = \begin{bmatrix} B_{11} & Y & X \\ Y^* & B_{22} & Z \\ X^* & Z^* & B_{33} \end{bmatrix}$$

which satisfies $\|B - A\| < \epsilon$, the matrix B_{11} is positive definite and B_{22} is negative definite.

Taking ϵ smaller, if necessary, and using Lemma 2.3, find $K > 0$ such that for any hermitian matrices B_{11} and B_{22} satisfying

$$\|I - B_{11}\|, \quad \|-I - B_{22}\| < \epsilon$$

there exist invertible matrices S_1 and S_2 with the properties that

$$\|I - S_1\| \leqq K\|I - B_{11}\|, \quad \|I - S_2\| \leqq K\|-I - B_{22}\|$$

and

$$S_1{}^*B_{11}S_1 = I, \quad -S_{22}{}^*B_{22}S_2 = I.$$

Then we can replace the B given by (2.5) by

$$\begin{bmatrix} S_1^* & 0 & 0 \\ 0 & S_2^* & 0 \\ 0 & 0 & I \end{bmatrix} B \begin{bmatrix} S_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & I \end{bmatrix}.$$

In other words, it is sufficient to prove Lemma 2.4 for B given by (2.5) with $B_{11} = I_m$ and $B_{22} = -I_p$. As, moreover, B is congruent to A, we actually have rank B = rank A, and hence

$$(2.6) \qquad B_{33} = [X^*Z^*]\begin{bmatrix} I & Y \\ Y^* & -I \end{bmatrix}^{-1}\begin{bmatrix} X \\ Z \end{bmatrix}.$$

Indeed, write

$$(2.7) \qquad \begin{bmatrix} I & Y & X \\ Y^* & -I & Z \\ X^* & Z^* & B_{33} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0.$$

Then

$$\begin{bmatrix} I & Y & x \\ Y^* & -I & y \end{bmatrix} + \begin{bmatrix} X \\ Z \end{bmatrix} z = 0,$$

$$[X^*Z^*]\begin{bmatrix} x \\ y \end{bmatrix} + B_{33}z = 0,$$

so

$$(2.8) \qquad \left(B_{33} - [X^*Z^*]\begin{bmatrix} I & Y \\ Y^* & -I \end{bmatrix}^{-1}\begin{bmatrix} X \\ Z \end{bmatrix} z\right) = 0.$$

But rank $B = m + p$ implies that equation (2.7) has $n - m - p$ linearly independent solutions which means that every $z \in \mathbf{C}^{n-m-p}$ satisfies (2.8), so (2.6) holds.

Given B as in (2.5) with B_{33} satisfying (2.6), we shall look for an invertible matrix T such that $T^*AT = B$, in the form

$$T^* = \begin{bmatrix} I & 0 & 0 \\ T_1 & T_2 & 0 \\ T_3 & T_4 & I \end{bmatrix}.$$

So

$$(2.9) \quad \begin{bmatrix} I & 0 & 0 \\ T_1 & T_2 & 0 \\ T_3 & T_4 & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I & T_1^* & T_3^* \\ 0 & T_2^* & T_4^* \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} I & Y & X \\ Y^* & -I & Z \\ X^* & Z^* & B_{33} \end{bmatrix}.$$

It is not difficult to see (by using (2.6) and the equality

$$X^*X - (X^*Y - Z^*)(I + Y^*Y)^{-1}(Y^*X - Z)$$
$$= [X^*Z^*] \begin{bmatrix} I & Y \\ Y^* & -I \end{bmatrix}^{-1} \begin{bmatrix} X \\ Z \end{bmatrix} \Bigg)$$

that (2.9) is satisfied with

$$(2.10) \quad T_1 = Y^*, \quad T_2 = (I + Y^*Y)^{1/2}, \quad T_3 = X^*,$$
$$T_4^* = (1 + Y^*Y)^{-1/2}(Y^*X - Z).$$

Lemma 2.4 obviously follows from (2.10).

We need one more lemma for the proof of Theorem 2.1.

LEMMA 2.5. *Let $A \neq 0$ be an $n \times n$ hermitian matrix. Then there exists $M > 0$ such that for every hermitian B which is congruent to A there exists an invertible S such that $S^*AS = B$ and*

$$(2.11) \quad \|S\|^2 \leqq M\frac{\|B\|}{\|A\|}.$$

Proof. We can assume that A is diagonal with

$$A = \operatorname{diag}(a_1, \ldots, a_m, a_{m+1}, \ldots, a_r, 0, \ldots, 0),$$

where $a_i > 0$ for $i = 1, \ldots, m$; $a_i < 0$ for $i = m + 1, \ldots, r$. Replacing B by U^*BU with unitary U, we can also assume that

$$b = \operatorname{diag}(b_1, \ldots, b_m, b_{m+1}, \ldots, b_r, 0, \ldots, 0),$$

where $b_i > 0$ for $i = 1, \ldots, m$; $b_i < 0$ for $i = m + 1, \ldots, r$. Take

$$S = \operatorname{diag}(s_1, \ldots, s_m, s_{m+1}, \ldots, s_r, s_{r+1}, \ldots, s_n), \quad s_i > 0,$$

where

$$s_i^2 = |b_i|/|a_i|, \quad i = 1, \ldots, r \quad \text{and}$$
$$s_i < \max(s_1, \ldots, s_r) \quad \text{for } i = r + 1, \ldots, n.$$

Then (2.11) holds with

$$M = \max(|a_1|^{-1}, \ldots, |a_r|^{-1})[\min(|a_1|, \ldots, |a_r|)]^{-1}.$$

Finally we are ready to prove Theorem 2.1. Given an $n \times n$ hermitian matrix A, by Lemma 2.4 there exist positive constants ϵ and K_0 such that

$$\Omega(A, B) \leqq K_0\|B - A\|$$

for any matrix B congruent to A such that $\|B - A\| < \epsilon$. Assume now that B is congruent to A and $\|B - A\| \geqq \epsilon$. By Lemma 2.5,

$$(2.11) \quad \|S\|^2 \leqq M\frac{\|B\|}{\|A\|}$$

for some invertible S such that $S^*AS = B$. Then

$$\|I - S\| \leqq 1 + \|S\| \leqq 1 + M\frac{(\|B\|)^{1/2}}{(\|A\|)^{1/2}}$$
$$\leqq 1 + M\left(\frac{\|B - A\|}{\|A\|} + 1\right)^{1/2}.$$

Let

$$L = \max\left(\epsilon^{-1}\left(1 + M\left(\frac{\epsilon}{\|A\|} + 1\right)^{1/2}\right), 1 + \frac{1}{2}\frac{M}{\|A\|}\right).$$

Then (provided $\|B - A\| \geqq \epsilon$)

$$1 + M\left(\frac{\|B - A\|}{\|A\|} + 1\right)^{1/2} \leqq L\|B - A\|,$$

and we obtain Theorem 2.1 with $K = \max(K_0, L)$.

The proof of Theorem 2.1 shows that Theorem 2.2 can be proved analogously with the help of the following result. We first recall that S_α is the set of all $n \times n$ hermitian matrices A such that $\kappa(A) \leqq \alpha$.

LEMMA 2.6. *For every $\alpha > 0$ there is a positive constant M such that any hermitian matrix $A \in S_\alpha$ with m positive and r negative eigenvalues (counting multiplicities) is congruent to*

$$\begin{bmatrix} I_m & 0 & 0 \\ 0 & -I_r & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

with the congruence matrix S satisfying

$$\|S\|, \|S^{-1}\| \leqq M.$$

The proof of Lemma 2.6 is immediate if one reduces A by a unitary similarity to a diagonal matrix.

Remark 2.7. If we restrict the matrices A and B to be real symmetric, all of the results of this section are still valid (here we define $\Omega(A, B) = \inf\{ \|I - S\| \,|S^*AS = B, S \text{ real, invertible}\} $).

3. Examples. In connection with Theorem 2.1 and Lemma 2.4 one might think that given a hermitian matrix A there is a positive constant K such that

$$\Omega(B, A) \leqq K\|B - A\| \tag{3.1}$$

for every B which is congruent to A. However, this is not so (in general) as the following example shows.

Example 3.1. Let $A = I$ and $B_\alpha = \alpha I$, $\alpha > 0$. Then $S \in \Omega(B, A)$ if and only if $\alpha S^*S = I$, or $\sqrt{\alpha}S$ is unitary. Hence $\|S\| = 1/\sqrt{\alpha}$ and (3.1) would imply (for $\alpha < 1$) that

$$1/\sqrt{\alpha} \leqq K|\alpha - 1|,$$

which is contradictory when $\alpha \to 0$.

The second example is of illustrative character. Here we compute explicitly $\Omega(A, B)$ in a particular 2×2 case.

Example 3.2. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

One verifies easily that $S^*AS = B$ holds if and only if

$$S = \begin{bmatrix} 0 & b \\ c & d \end{bmatrix}$$

with $|b| = 1$. Using the fact that $\|I - S\|^2$ coincides with the largest eigenvalue of the matrix $(I - S^*)(I - S)$, we find that

$$\|I - S\|^2 = (1/2)[2 + |c|^2 + |1 - d|^2 + \sqrt{u}], \tag{3.2}$$

where

$$u = 2 + |c|^2 + |1 - d|^2 + 4|b + \overline{c} - \overline{c}d|^2.$$

A calculation shows that

$$2\Omega(A, B)^2 = \inf[2 + x^2 + y^2 + \sqrt{v}], \tag{3.3}$$

where

$$v = (2 + x^2 + y^2)^2 + 4(1 - xy)^2,$$

and the infimum is taken over all $x > 0$ and $y \geqq 0$. If $x^2 + y^2$ is fixed and is $\geqq 2$, then the infimum in the right hand side of (3.3) is achieved when $xy = 1$. If $x^2 + y^2$ is fixed and is <2, then the infimum in the right hand side of (3.3) is achieved when $x = y$. It follows that the infimum in (3.3) over the set $\{x > 0, y \geqq 0\}$ is obtained when $x = y$ and $x \to 0$ which gives

$$\Omega(A, B)^2 = 1 + \sqrt{2}.$$

This example shows that in general the infimum in the definition of

$$\Omega(A, B) = \inf\{ \|I - S\| \,|S^*AS = B, S \text{ invertible}\}$$

is not attained, i.e., there is no invertible S such that

$$(3.4) \qquad S^*AS = B \quad \text{and} \quad \Omega(A, B) = \|I - S\|.$$

It is not difficult to see that in case A is an invertible hermitian matrix and B is congruent to A, then there is an invertible S with the properties (3.4). Indeed, let $\{S_m\}$, $m = 1, 2, \dots$ be a sequence of invertible matrices such that

$$S_m{}^*AS_m = B \quad \text{and} \quad \|I - S\| \to \Omega(A, B).$$

This sequence is obviously bounded, so we can assume that $S_m \to S$ for some S. As $S^*AS = B$, and B was assumed to be invertible, S is invertible as well, and $\|I - S\| = \Omega(A, B)$.

Finally, let us demonstrate that in general $\Omega(B, A) \neq \Omega(A, B)$.

Example 3.3. Let $A = I$ and $B = \alpha I$, where $\alpha > 0$. Then $S^*AS = B$ if and only if $S = \sqrt{\alpha}U$ for some unitary matrix U. Thus

$$\|I - S\| = \|U^{-1} - \sqrt{\alpha}I\|$$

and

$$\Omega(A, B) = \inf\|W - \sqrt{\alpha}I\|,$$

where the infimum is taken over all unitary matrices W. Reducing W to a diagonal form by a unitary similarity, we see easily that

$$\Omega(A, B) = |1 - \sqrt{\alpha}|.$$

An analogous consideration shows that

$$\Omega(B, A) = |1 - 1/\sqrt{\alpha}|.$$

Thus unless $\alpha = 1$, $\Omega(A, B) \neq \Omega(B, A)$.

4. Unitary congruence. Here we state and prove results analogous to Theorems 2.1 and 2.2 for the case of congruence by unitary matrices. We shall do that in the framework of normal matrices rather than hermitian matrices (this is a more natural framework in which to consider unitary congruence, and the proofs are not affected by this extension).

THEOREM 4.1. *Let A be a normal $n \times n$ matrix. Then there is a constant $K > 0$ such that for every normal matrix B which is unitarily congruent to A there is a unitary matrix U such that $B = U^*AU$ and*

$$\|I - U\| \leqq K\|B - A\|.$$

We shall actually prove the following more general result. For every $n \times n$ matrix $A \neq 0$, define

$$\mu(A) = \max\{ |\lambda_i - \lambda_j|, |\lambda_i - \lambda_j|^{-1}, i \neq j, i, j = 1, \dots, p\},$$

where $\lambda_1, \dots, \lambda_p$ are all the distinct eigenvalues of A. For every $\alpha > 0$, let T_α be the set of all $n \times n$ nonzero matrices A such that $\mu(A) \leqq \alpha$.

THEOREM 4.2. *Let $\alpha > 0$ be given. Then there is a $K > 0$ such that for every normal $A \in T_\alpha$ and every B which is unitarily congruent to A there is a unitary matrix U such that $B = U^*AU$ and*

$$\|I - U\| \leqq K\|B - A\|. \tag{4.1}$$

In the proof of Theorem 4.2 we shall use continuity properties of the Moore-Penrose generalized inverse.

Recall that an $m \times n$ matrix B is called the *Moore-Penrose generalized inverse* of an $n \times m$ matrix A if $BAB = B$, $ABA = A$, and the matrices AB and BA are hermitian (see, e.g., [**1**] for information on generalized inverses). The Moore-Penrose generalized inverse of A is unique and will be denoted A^I. Observe that AA^I is the orthogonal projection on the range of A and A^IA is the orthogonal projection on the range of A^I. Also, the ranks of A and A^I coincide.

PROPOSITION 4.3. *Let $R_{m,n}(s)$ be the set of all $m \times n$ complex matrices with fixed rank s. Then the function*

$$f: R_{m,n}(s) \to R_{n,m}(s)$$

defined by $f(A) = A^I$ is continuous. Moreover, for every $\alpha > 0$ there exists a positive constant K such that

$$\|A^I - B^I\| \leqq K\|A - B\| \tag{4.2}$$

for every pair of matrices $A, B \in R_{m,n}(s)$ satisfying $\|A^I\|, \|B^I\| \leqq \alpha$.

Proof. This proposition follows from Theorems 10.4.3 and 10.4.5 in [**1**].

Proof of Theorem 4.2. Since $\|I - U\| \leqq 2$ for every unitary matrix U, it is sufficient to prove (4.1) for all normal matrices B which are unitarily similar to A and such that $\|B - A\| < \epsilon$ (where $\epsilon > 0$ depends only on α). We may assume that

$$A = \operatorname{diag}(\lambda_1 I_{k_1}, \ldots, \lambda_r I_{k_r}),$$

where $\lambda_1, \ldots, \lambda_r$ are distinct complex numbers. For any B which is unitarily similar to A, we have

$$\dim \operatorname{Ker}(\lambda_i - A) = \dim \operatorname{Ker}(\lambda_i - B), \quad i = 1, \ldots, r.$$

Thus, applying Proposition 4.3 and using the fact that

$$\|\lambda_i - B\| = \max_{j \neq i} |\lambda_j - \lambda_i|, \quad \text{and}$$
$$\|(\lambda_i - B)^I\| = \max_{j \neq i} |\lambda_j - \lambda_i|^{-1},$$

we obtain

$$(4.3) \quad \|[I - (\lambda_i - B)^I(\lambda_i - B)] - [I - (\lambda_i - A)^I(\lambda_i - A)]\| \leqq K_1 \|B - A\|,$$

where K_1 depends only on α.

Observe that $Q_i(B) = I - (\lambda_i - B)^I(\lambda_i - B)$ is by definition the orthogonal projection on $\operatorname{Ker}(\lambda_i - B)$. Let

$$x_{i1}, \ldots, x_{ik_i}$$

be an orthonormal basis of $\operatorname{Ker}(\lambda_i - A)$ and let

$$y_{i1}, \ldots, y_{ik_i}$$

be the vectors obtained by applying the Gram-Schmidt process to the set of k_i vectors

$$Q_i(B)x_{ij}, \quad j = 1, \ldots, k_i.$$

Because of (4.3), for B close enough to A, $\{y_{ij}, j = 1, \ldots, k_i\}$ form an orthonormal basis for $\operatorname{Ker}(\lambda_i - B)$ and

$$\|x_{ij} - y_{ij}\| \leqq K_1 \|B - A\|$$

for all i and j, where K_1 depends only on α. Let U be defined by the properties that $Uy_{ij} = x_{ij}$, $j = 1, \ldots, k_i$, $i = 1, \ldots, r$. Then $B = U^*AU$ and (4.1) holds.

5. Simultaneous congruence. We now fix r $n \times n$ hermitian matrices $A_1, \ldots, A_r$. We say that $A_1, \ldots, A_r$ are *simultaneously congruent* to $B_1, \ldots, B_r$ if there exists an invertible matrix S such that

$$S^*A_iS = B_i, \quad i = 1, \ldots, r.$$

If $A_1, \ldots, A_r$ are simultaneously congruent to $B_1, \ldots, B_r$, define

$$\Omega(A_1, \ldots, A_r; B_1, \ldots, B_r) = \inf \|I - S\|$$

where the infimum is taken over all invertible S such that $S^*A_iS = B_i$, $i = 1, \ldots, r$.

In this section, we compare $\Omega(A_1, \ldots, A_r; B_1, \ldots, B_r)$ with

$$\sum_{i=1}^{r} \|B_i - A_i\|.$$

The following example shows that a result analogous to Theorem 2.1 is false in the general situation.

Example 5.1. Let

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

We assert that

$$\frac{\Omega(A_1, A_2; B_1, B_2)}{\|B_1 - A_1\| + \|B_2 - A_2\|}$$

is unbounded as (B_1, B_2) runs over all pairs simultaneously congruent to (A_1, A_2).

Choose a positive number α and let

$$S_\alpha = \begin{bmatrix} 0 & \alpha^{-1} \\ \alpha & 0 \end{bmatrix}.$$

Then $S_\alpha^* A_1 S_\alpha = A_1$ and

$$S_\alpha^* A_2 S_\alpha = \begin{bmatrix} \alpha^2 & 1 \\ 1 & 0 \end{bmatrix} = B_{2,\alpha}.$$

Suppose T is any other matrix such that

$$T^* A_1 T = A_1 \quad \text{and} \quad T^* A_2 T = B_{2,\alpha}.$$

It is easy to check that T must have the form

$$T = \begin{bmatrix} r & s \\ t & 0 \end{bmatrix}$$

with $|s| = \alpha^{-1}$, $|t| = \alpha$, $s\bar{t} = 1$, and $\operatorname{Re}(r\bar{t}) = 0$.

As $\alpha \to 0$, $\|B_{2,\alpha} - A_2\| \to 1$, but $\|I - T\| \geqq |\alpha|^2 + |\alpha|^{-2} \to \infty$. This verifies Example 5.1.

However, we propose the following conjecture.

CONJECTURE 5.2. *Let $A_1, \ldots, A_r$ be as above. In addition, assume that A_1 is positive definite. Then there is a positive constant K depending only on $A_1, \ldots, A_r$ such that if S is invertible and $B_i = S^* A_i S$, $i = 1, \ldots, r$, then*

$$\Omega(A_1, \ldots, A_r; B_1, \ldots, B_r) \leqq K \sum_{i=1}^{r} \|B_i - A_i\|.$$

We believe that the conjecture is true and that the compactness of the group of all invertible S such that $S^*A_iS = A_i$, $i = 1, \ldots, r$ should be crucial for its proof.

Finally, we remark that the affirmative solution of Conjecture 5.2 allows one to obtain the following statement concerning simultaneous unitary similarity:

Let $A_1, \ldots, A_r$ be hermitian matrices, and for every r-tuple $B_1, \ldots, B_r$ of hermitian matrices such that

$$(5.1) \qquad B_j = U^*A_jU, \quad j = 1, \ldots, r,$$

for some unitary matrix U, denote

$$\Omega(A_1, \ldots, A_r; B_1, \ldots, B_r) = \inf\|I - U\|$$

where the infimum is taken over all unitary U's with property (5.1). Then there exists a positive constant K (depending on $A_1, \ldots, A_r$ only) such that

$$\Omega(A_1, \ldots, A_r; B_1, \ldots, B_r) \leqq K \sum_{i=1}^{r} \|B_i - A_i\|$$

for every r-tuple $B_1, \ldots, B_r$ which is simultaneously unitarily similar to $A_1, \ldots, A_r$. Indeed, apply Conjecture 5.2 to the $(r + 1)$-tuple $(I, A_1, \ldots, A_r)$.

References

1. S. L. Campbell and C. D. Meyer, Jr., *Generalized inverses of linear transformations* (Pitman, London, etc., 1979).

2. I. Gohberg and L. Rodman, *Perturbations of lattices of invariant subspaces of matrices*, Lin. Alg. and Appl. *76* (1985), 85-120.

3. I. Gohberg, P. Lancaster and L. Rodman, *Invariant subspaces of matrices with applications* (J. Wiley, 1986).

San Diego State University,
San Diego, California;
Arizona State University,
Tempe, Arizona