Two Theorems on Matrices*

Morris Newman

(May 16, 1962)

Generalizations of theorems important in the iterative solution of systems of linear equations are given, together with a lemma on the solution of a certain matrix equation.

References

In this note we generalize two theorems on matrices which are of importance in certain iterative schemes for the solution of systems of linear equations (see [1] and [3] for relevant references).

The first well-known and the result is supported by the Office of Naval Research. Figures in brackets indicate the literature references at the end of this paper.

THEOREM 1. If R and S are both p.d. or both n.d. then every eigenvalue of M is of modulus less than 1. If R is p.d. and S n.d. or if R is n.d. and S p.d. then every eigenvalue of M is of modulus greater than 1.

THEOREM 2. If every eigenvalue of M is of modulus less than 1, then R is p.d. if S is p.d. and n.d. if S is n.d. If every eigenvalue of M is of modulus >1, then R is p.d. if S is n.d. and n.d. if S is p.d.

Before giving the proofs of these theorems we state two lemmas. The first is well-known and we omit the proof. The proof of the second is given in detail.

Let $A = (a_{ij})$ be an arbitrary $n \times n$ complex matrix. If $B = (b_{ij})$ is a non-negative matrix such that $a_{ij} \leq b_{ij}$, we shall write $A \preceq B$. It is trivial to verify that if $A_1 \preceq B_1$, $A_2 \preceq B_2$ then $A_1 + A_2 \preceq B_1 + B_2$, $A_1A_2 \preceq B_1B_2$. Further, let $M(A)$ denote the absolute value of an eigenvalue of largest modulus of $A$.

LEMMA 1. Let $A$ be an arbitrary $n \times n$ complex matrix. Then there is a fixed non-negative matrix $B$ such that for every integer $r \geq 0$,

$$A^r \preceq r^{n-1}M(A)B.$$  (2)

This lemma is well-known and a proof can be given using the Jordan normal form of the matrix $A$. The term $r^{n-1}$ can be replaced by $r^{m-1}$, where $m$ is the order of a largest Jordan block of $A$. See also the paper [2] by A. Ostrowski.

LEMMA 2. Let $A, B, C$ be $n \times n$ complex matrices and suppose that

$$\rho = M(A)M(B) < 1.$$  (3)

Then the matrix equation

$$X - AXB = C$$  (4)

has a unique solution $X$ given by the infinite series

$$X = \sum_{r=0}^{\infty} A^rCB^r = C + ABC + A^2CB^2 + \ldots.$$  (5)

PROOF. Set

$$X_0 = C, X_{r+1} = AX_rB + C, r \geq 0.$$  (6)

so that

$$X_r = C + ABC + \ldots + A^rCB^r.$$  (7)

If $r$ and $p$ are non-negative integers then it is easy to verify that

$$X_{r+p} - X_{r-1} = A^rX_pB^r.$$  (8)

By lemma 1 there are non-negative matrices $U$, $V$, $W$, independent of $r$ such that

$$A^r \preceq r^{n-1}M(A)U, C \preceq V, B^r \preceq r^{n-1}M(B)W.$$  (9)

Set $T = UVW$. Then

$$X_p = \sum_{r=0}^{p} A^rCB^r \preceq \left(\sum_{r=0}^{p} r^{n-2}\rho^r\right)T,$$

where $\rho$ is defined by (3). Since $0 \leq \rho < 1$, this implies that $\sum_{r=0}^{\infty} r^{n-2}\rho^r$ converges, and that

$$X_p \preceq \left(\sum_{r=0}^{\infty} r^{n-2}\rho^r\right)T.$$  (10)

Hence lemma 1 implies that there is a non-negative matrix $K$ independent of $r$ and $p$ such that

$$X_{r+p} - X_{r-1} \preceq r^{n-2}\rho^rK.$$  (11)

From (3) and (11) it follows easily that the sequence $\{X_r\}$ is a Cauchy sequence converging to a limit $X$, and (6) implies that $X = AXB + C$ and that $X$ satisfies (5). To show that the solution $X$ is
unique, let \( Y \) be any other solution. Then \( Y - X = A(Y - X)B \) and by iteration, \( Y - X = A^r(Y - X)B^r \) for every non-negative integer \( r \). Hence there is a non-negative matrix \( L \) independent of \( r \) such that \( Y - X < r^{2-n} \rho_r L \). Letting \( r \) go to infinity, (3) implies that \( Y - X = A(Y - X)B \) and by iteration, \( Y - X = A^r(Y - X)B^r \) for every non-negative integer \( r \). Hence there is a non-negative matrix \( L \) independent of \( r \) such that \( Y - X < 1/2^p \). Letting \( r \) go to infinity, (3) implies that \( Y = X \). The proof of the lemma is complete. We go on to the proof of theorem 1.

Let \( \lambda \) be an eigenvalue of \( M \), \( \nu \) a corresponding eigenvector, and set \( f = \nu^* F \nu, \quad r = \nu^* R \nu, \quad s = \nu^* S \nu \). Then (1) implies that

\[
 r = f + \lambda f, \quad s = f - \lambda f. \tag{8}
\]

Assume that \( R \) and \( S \) are definite matrices. Then \( \lambda \neq -1 \) and we can eliminate \( f \) from (8). We obtain

\[
 r = \frac{1 - |\lambda|^2}{s} \tag{9}
\]

Theorem 1 is implied by (9) since \( r/s > 0 \) if \( R \) and \( S \) are both p.d. or both n.d., and \( r/s < 0 \) if \( R \) is p.d. and \( S \) n.d. or \( R \) is n.d. and \( S \) p.d.

We now turn to the proof of theorem 2. If we eliminate \( F \) from (1) (remembering that \( R \) and \( S \) are hermitian) we obtain

\[
(I + M^*) S (I + M) = R - M^* R M. \tag{10}
\]

If we assume that every eigenvalue of \( M \) is of modulus less than 1 then lemma 2 applies and we obtain

\[
 R = \sum_{r=0}^{\infty} (M^*)^r (I + M^*) S (I + M) M^r. \tag{11}
\]

Since the sum of any number of p.s.d. matrices is p.s.d. and the sum of any number of n.s.d. matrices n.s.d., the matrix \( R \) is p.s.d. if \( S \) is p.s.d. and n.s.d. if \( S \) is n.s.d. But the first term of the sum (11) is p.d. if \( S \) is p.d. and n.d. if \( S \) is n.d. Thus the first part of theorem 2 is proved.

Assume now that every eigenvalue of \( M \) is of modulus greater than 1. Then \( M \) is nonsingular and every eigenvalue of \( N = M^{-1} \) is of modulus less than 1. Equation (10) may be written

\[
(I + N^*) S (I + N) = N^* R N - N. \tag{12}
\]

Once again lemma 2 applies and we obtain

\[
 R = -\sum_{r=0}^{\infty} (N^*)^r (I + N^*) S (I + N) N^r, \tag{13}
\]

from which the second part of theorem 2 follows. This concludes the proof of theorem 2.

For related theorems, see the paper [4] by P. Stein. Lemma 2 of our paper can be used to provide a simplified proof of the theorems given by Stein in [4] and indeed to generalize these theorems.

References


(Paper 66B3–76)