JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 66B, No. 3, July–September 1962

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).[1] In addition we prove a lemma on the solution of certain matrix equations. Although the proofs are simple, the results are of interest and find application in computations with matrices. However, we forego the discussion of these applications here.

In what follows F and M are $n \times n$ matrices and "*" denotes "conjugate transpose." We set

$$R=F+FM,\ S=F^*-FM \tag{1}$$

and assume that R and S are hermitian. We write p.d. for positive definite, n.s.d. for negative semi-definite, etc. The theorems we wish to prove follow.

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 $A<<B$. It is trivial to verify that if $A_1<<B_1$, $A_2<<B_2$ then $A_1+A_2<<B_1+B_2$, $A_1A_2<<B_1B_2$. 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* $r\geq 0$,

$$A^r<<r^{n-1}M(A)^rB. \tag{2}$$

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.

*The preparation of this paper was supported by the Office of Naval Research.
[1] Figures in brackets indicate the literature references at the end of this paper.

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

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

Then the matrix equation

$$X-AXB=C \tag{4}$$

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

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

PROOF. Set

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

so that

$$X_r=C+ACB+\ldots+A^rCB^r.$$

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

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

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

$$A^r<<r^{n-1}M(A)^rU,\ C<<V,\ B^r<<r^{n-1}M(B)^rW.$$

Set $T=UVW$. Then

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

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

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

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

$$X_{r+p}-X_{r-1}<<r^{2n-2}\rho^rK. \tag{7}$$

From (3) and (7) 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 $Y-X<<r^{2n-2}\rho^r L$. 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 λ be an eigenvalue of M, v a corresponding eigenvector, and set $f=v^*Fv$, $r=v^*Rv$, $s=v^*Sv$. Then (1) implies that

$$r=f+\lambda f,\ s=\bar{f}-\lambda f. \tag{8}$$

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

$$\frac{r}{s}=\frac{1-|\lambda|^2}{|1+\lambda|^2} \tag{9}$$

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

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

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

$$(I+N^*)S(I+N)=N^*RN-N. \tag{12}$$

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

[1] F. John, Advanced numerical analysis, New York University (1956).

[2] A. Ostrowski, Über Normen von Matrizen, Math. Zt. **63**, 2–18 (1955).

[3] E. Reich, On the convergence of the classical iterative method of solving linear simultaneous equations, Ann. Math. Stat. **20**, 448–451 (1949).

[4] P. Stein, Some general theorems on iterants, J. Research NBS **48**, 82–83 (1952).

(Paper 66B3–76)