Journal of Research of the National Bureau of Standards Vol. 48, No. 1, January 1952 Research Paper 2288

Some General Theorems on Iterants [1]

P. Stein [2]

If B is a square matrix, then it is known that a necessary and sufficient condition that $\lim_{n\to\infty} B^n=0$, is that the characteristic roots of B are all of modulus less than unity. An alternative condition is given in this paper, in terms of Hermitian matrices. Further, a generalization of the result is obtained that covers cases of matrices B whether B^n does or does not converge to 0, except for very special matrices.

Introduction. If B is a square matrix with real or complex elements, it is well known that a necessary and sufficient condition that $\lim_{n\to\infty} B^n=0$ is that the characteristic roots of B are all of modulus less than 1.

In this paper an alternative condition for the convergence of B^n to 0 will be given in terms of certain Hermitian and symmetric matrices. We also obtain a generalization of this result that covers matrices B when B^n does or does not converge to 0, except for a special class of such matrices B.

My thanks are due to Olga Taussky Todd for the help and encouragement I received from her in the preparation of this paper. My thanks are also due to L. J. Paige for valuable suggestions toward the improvement and simplification of the proofs.

We will consider square matrices B whose elements are either real or complex. The conjugate transpose of B will be denoted by B^*.

THEOREM 1. A necessary and sufficient condition that $\lim_{n\to\infty} B^n=0$ is that there exist a positive definite Hermitian matrix H for which $H-B^*HB$ is positive definite.

Corollary 1. If B is real, H may be taken real and symmetric.

Proof: Necessity: Let P be a nonsingular matrix such that

$$PBP^{-1}=K_1\dot{+}K_2\dot{+}\ \ldots\ \dot{+}K_r,$$

where K_i is the Jordan normal form; i. e.

$$K_i=\lambda_i I^{n_i\times n_i}+U^{n_i\times n_i},$$

where $\sum_{i=1}^{r} n_i=n$, λ_i are the *not necessarily distinct* characteristic roots of B, and $U^{n_i\times n_i}$ is a matrix with units in the superdiagonal and zero elsewhere.

Let $\delta_i=\delta(\epsilon_i)$ be the diagonal matrix $(\epsilon_i^{n_i-1}, \epsilon_i^{n_i-2}, \ldots, 1)$ for $i=1, 2, \ldots r$.

If $Q=\delta_1\dot{+}\delta_2\dot{+}\ldots\dot{+}\delta_r$, then

$$K=QPBP^{-1}Q^{-1}=N_1\dot{+}N_2\dot{+}\ \ldots\ \dot{+}N_r, \text{ where}$$

$$N_i=\delta_i K_i\delta_i^{-1}=\lambda_i I+\epsilon_i U;$$

it being understood that I and U are of the correct order.

Note that

$$I-K^*K=(I-N_1^*N_1)\dot{+}(I-N_2^*N_2)\dot{+}\ \ldots \\ \dot{+}(I-N_r^*N_r) \tag{2}$$

is positive definite if and only if $(I-N_i^*N_i)$ is positive definite for all i. Clearly

$$I-N_i^*N_i=(1-\bar{\lambda}_i\lambda_i)I-\epsilon_i(\bar{\lambda}_iU+\lambda_iU^*+\epsilon_iU^*U) \tag{3}$$

will be positive definite if

$$1-\bar{\lambda}_i\lambda_i>\epsilon_i\left[\frac{y^*(\bar{\lambda}_iU+\lambda_iU^*)y}{y^*y}+\epsilon_i\frac{y^*U^*Uy}{y^*y}\right] \text{for } y\neq 0. \tag{4}$$

If $M=\max|\lambda_i|$, we have

$$\frac{y^*(\bar{\lambda}_iU+\lambda_iU^*)y}{y^*y}<2M, \text{ also } \frac{y^*U^*Uy}{y^*y}<1 \text{ for all } y\neq 0;$$

hence

$$\epsilon_i\left[\frac{y^*(\bar{\lambda}_iU+\lambda_iU^*)y}{y^*y}+\epsilon_i\frac{y^*U^*Uy}{y^*y}\right]< \\ \epsilon_iM+\epsilon_i^2 \text{ for all } y\neq 0. \tag{5}$$

Since $\lim_{n\to\infty} B^n=0$, $|\lambda_i|<1$; hence, from (3), (4), and (5), $I-N_i^*N_i$ is positive definite for sufficiently small values of ϵ_i; and so from (2), $I-K^*K$ is positive definite for such values of ϵ_i. A change of variable $y=QPx$ gives

$$y^*(I-K^*K)y=x^*(H-B^*HB)x, \text{ where } H=P^*Q^*QP.$$

Since H is clearly positive definite, the proof for the necessity part is complete.

Sufficiency:[3] Let H be *any* positive definite Hermitian matrix for which $H-B^*HB$ is positive definite. Since H is positive definite, $H=D^*D$, and by making the change of variables $Dx=y$,

$$x^*(H-B^*HB)x=y^*(I-K^*K)y>0, \qquad (K=DBD^{-1}). \tag{6}$$

[1] This work was performed under a National Bureau of Standards contract with the University of California at Los Angeles, and was sponsored in part by the Office of Naval Research.

[2] University of Natal, Durban, South Africa, and University of California at Los Angeles.

[3] This proof was suggested by L. J. Paige.

Now, if λ_i is any characteristic root of K (and hence of B), y_i an associated characteristic vector, we see that

$$y_i^*(I-K^*K)y_i=y_i^*y_i-\bar{\lambda}_i\lambda_i y_i^*y_i>0.$$

Thus, $|\lambda_i|<1$ for all characteristic roots of B, and hence B^n will converge to 0.

To prove the corollary, we suppose the elements of B real. Let H be the matrix of the theorem, then $H=A+iS$, where A is a real symmetric matrix, and S is a real skew-symmetric matrix. If H is positive definite, then it is known that A is positive definite. Again

$$H-B^*HB=H-B'HB=A-B'AB+i\,(S-B'SB).$$

$A-B'AB$ is symmetric and $S-B'SB$ is skew-symmetric. If $H-B'HB$ is positive definite, then $A-B'AB$ is positive definite. Hence we may use A in place of H in the theorem.

We give a sufficiency test for the nonconvergence of B^n to 0.

Theorem 2. If there exists a nonpositive-definite matrix H such that $H-B^*HB$ is positive definite, then $\lim_{n\to\infty} B^n\neq 0$.

For proof, we observe that if H is not positive definite, a vector x may be found such that $x^*Hx\leq 0$. Further, if $H-B^*HB$ is positive definite, the sequence $x^*B^{*n}HB^nx$ is decreasing. Hence $\lim_{n\to\infty} B^nx\neq 0$ and so $\lim_{n\to\infty} B^n\neq 0$.

It may be observed that the condition that $H-B^*HB$ should be positive definite may be weakened to $H-B^*HB$ at least positive-semi-definite, provided H is not positive-semi-definite.

Now we shall prove a generalization of the necessity part of theorem 1.

Theorem 3. Let B be a matrix whose characteristic roots of modulus 1 have multiplicity no greater than two. Then there exists a nonzero Hermitian matrix H_1 such that $H_1-B^*H_1B\geq 0$.

Corollary 2. If B is real, H_1 may be taken real and symmetric. Proof. Using the expression (2) for $(I-K^*K)$, we see that

$$I-K^*\bar{K}-K^*(I-K^*K)K=[I-N_1^*N_1-N_1^*(I-N_1^*N_1)N_1]\dot{+}\ldots\dot{+}[I-N_r^*N_r-N_r^*(I-N_r^*N_r)N_r]$$

and again, this will be positive semidefinite if

$$(I-N_i^*N_i)-N_i^*(I-N_i^*N_i)N_i=(1-\bar{\lambda}_i\lambda_i)^2I-2\epsilon_i(1-\bar{\lambda}_i\lambda_i)E+\epsilon_i^2[\lambda_iU^*E+\bar{\lambda}_iEU+\epsilon_iU^*EU], \quad (8)$$

$\{E=(\bar{\lambda}_iU+\lambda_iU^*+\epsilon_iU^*U)\}$, is positive semidefinite.

Obviously, by a proper choice of ϵ_i, (8) can be made positive definite if B has no characteristic roots of modulus 1.

If B has a characteristic root such that $\bar{\lambda}_i\lambda_i=1$, the right side of (8) vanishes for roots of multiplicity 1. For roots of multiplicity two, (8) becomes $\begin{pmatrix}0 & 0\\ 0 & 2\epsilon_i^2\end{pmatrix}$ and hence can be made positive semidefinite.

Now a simple change of variables, $y=QPx$, as in Theorem 1, yields

$$y^*[(I-K^*K)-K^*(I-K^*K)K]y = x^*[(H-B^*HB)-B^*(H-B^*HB)B]x\geq 0,$$

where $H=P^*Q^*QP$. Thus the H_1 of our theorem is chosen as $H-B^*HB$.

If the multiplicity of a root of modulus 1 is three or greater, the right side of (8) is not positive semidefinite since it will always contain the principal subminor $\epsilon_i^2\begin{pmatrix}0 & \bar{\lambda}_r^2\\ \lambda_r^2 & \epsilon_i^2+2\bar{\lambda}_i\lambda_i\end{pmatrix}$.

Hence the method used in the proof of this theorem does not yield an H_1 in these cases.

It may be observed that $\lim_{n\to\infty} B^n=0$, if and only if H_1 is positive definite. For, if B has no roots of modulus equal to 1, then from the proof of the theorem it follows that $H_1-B^*H_1$B is positive definite, and the results follow from the sufficiency part of Theorem 1 and from Theorem 2. If B has a root of modulus 1, then since $H_1=H-B^*HB$, we may show, as in the proof of the sufficiency part of Theorem 1, that H_1 is at best positive semidefinite, and hence also not positive definite. In this case also $\lim_{n\to\infty} B^n\neq 0$.

Corollary 2 may be proved in the same way as corollary 1 of theorem 1.

Los Angeles, October 12, 1951.