Generalizations of Theorems of Lyapunov and Stein

Harald K. Wimmer
Technische Hochschule Graz
Graz, Austria

Recommended by H. Schneider

ABSTRACT

The matrix equation $f_H(A) = \sum c_{ij} A_i^{*} H A_j = W$, $H > 0$, $W > 0$, is studied. In the case $A^* H + H A = W$ [$H - A^* H A = W$], the controllability matrix of $(A^*, W)$ is used to determine the number of eigenvalues of $A$ on the imaginary axis [on the unit circle]. As an application a result of Pták on the critical exponent of the spectral norm is obtained. Estimates for the eigenvalues of $A$ satisfying $f_H(A) = M$ are given.

1. INTRODUCTION

In this note we deal with the matrix equation

$$\sum_{i,j=0}^{n-1} c_{ij} A_i^{*} H A_j = W, \quad c_{ij} = c_{ji}^{*},$$

where $A$ is a complex $n \times n$ matrix, $H$ and $W$ are hermitian, $H > 0$ (positive definite) and $W > 0$ (positive semidefinite). We focus (excepting the last section) on a semidefinite $W$. The case $W > 0$ has been studied by Hill [5] in the more general setting of the inequality

$$\sum_{i,j=1}^{s} c_{ij} A_i^{*} H A_j > 0$$

with quasi-commutative $A_1, \ldots, A_s$.

Special cases of (1) are the Lyapunov equation

$$A^* H + H A = W$$

and the Stein equation [13]
\[ H - A^*HA = W. \]  \hspace{1cm} (3)

Our results on (3) will be applied to determine the critical exponent of the spectral norm. In the last section we establish inequalities for the eigenvalues of a matrix A satisfying (1) or (2).

2. GENERAL THEOREMS

We define
\[ f_H(A) = \sum_{i,f=0}^{n-1} c_{ij}A^*iHA^i \]
and
\[ f(\lambda) = \sum_{i,f=0}^{n-1} c_{ij}\overline{\lambda}^i\lambda^j. \]  \hspace{1cm} (4)

As we assume \( c_{ij} = \overline{c_{ji}} \), \( f_H(A) \) is hermitian and \( f(\lambda) \) is real.

**Theorem 1.** [5,7]. If A satisfies (1) with \( H > 0 \), then
\( \begin{align*} \text{(a) } & W > 0, \quad (a') \ W \geq 0, \quad (a'') \ W = 0 \\
\text{(b) } & f(\lambda) > 0, \quad (b') \ f(\lambda) \geq 0, \quad (b'') \ f(\lambda) = 0 \end{align*} \)

implies
\( \begin{align*} \text{(b) } & f(\lambda) > 0, \quad (b') \ f(\lambda) \geq 0, \quad (b'') \ f(\lambda) = 0 \end{align*} \)

for each eigenvalue \( \lambda \) of A.

In order to determine the number of eigenvalues of A with \( f(\lambda) = 0 \) we introduce the following concepts.

**Definition** [11]. The pair \( (A,B) \), where A is \( n \times n \) and B is \( n \times m \), is called controllable if
\[ \text{rank}(B,AB,A^2B,\ldots,A^{n-1}B) = n. \]

The \( n \times nm \) matrix \( S(A|B) \), the controllability matrix of \( (A,B) \), is defined by
\[ S(A|B) = (B,AB,A^2B,\ldots,A^{n-1}B). \]
**THEOREMS OF LYAPUNOV AND STEIN**

**Lemma 1.** [4, 11]. \((A, B)\) is controllable if and only if

\[
\text{rank}(A - \lambda I, B) = n
\]

for each eigenvalue \(\lambda\) of \(A\).

**Lemma 2.** [12]. If \(V\) is hermitian and \(W = V^2\), then

\[
\text{rank} S(A|W) = \text{rank} S(A|V).
\]

**Theorem 2.** Let \(A\) satisfy (1) with \(H > 0\) and \(W > 0\). The following statements are equivalent:

(a) \(f(\lambda > 0\) for each eigenvalue \(\lambda\) of \(A\).

(b) The pair \((A^*, W)\) is controllable.

**Proof.** \(\gamma (b) \Rightarrow \gamma (a)\). Suppose \((A^*, W)\) or—in Lemma 2—\((A^*, V)\) is not controllable; then by Lemma 1 there is a \(u \neq 0\) and a \(\kappa\) such that \(u^*(A^* - \kappa I) = 0\) and \(u^* V = 0\). Then \(V f_H(A)u = f(\kappa) u^* Hu = u^* W u = 0\), and from \(H > 0\) we get \(f(\kappa) = 0\). By similar arguments we can show \(\gamma (a) \Rightarrow \gamma (b)\).

Given an additional condition on \(A\) Theorem 2 can be refined.

**Theorem 3.** Let \(r > 0\) be the number of (not necessarily distinct) eigenvalues \(\lambda\) of \(A\) with \(f(\lambda) = 0\), and suppose that the elementary divisors corresponding to these eigenvalues are all linear. If \(A\) satisfies (1) with \(H > 0\) and \(W > 0\), then

\[
r = n - \text{rank} S(A^*|W).
\]

**Proof.** Put \(\hat{A} = H^{1/2} A H^{-1/2}\) and \(\hat{W} = H^{-1/2} W H^{-1/2}\); then (1) is equivalent to \(f_1(\hat{A}) = \hat{W}\), and \(\text{rank} S(A^*|W) = \text{rank} S(\hat{A}^*|\hat{W})\). Thus without loss of generality we can assume \(H = I\) in (1).

Let \(\lambda_1, \ldots, \lambda_r\) be the \(r\) eigenvalues of \(A\) with \(f(\lambda) = 0\), and let \(u_1, \ldots, u_r\) be the corresponding eigenvectors, which exist by assumption. Then

\[
u_p^* f_1(A) u_p = u_p^* u_p f(\lambda_p) = u_p^* W u_p = 0
\]

and

\[
W u_p = 0, \quad (A - \lambda_p I) u_p = 0.
\]

The \(u_p\)'s are eigenvectors of the hermitian matrix \(W\); hence they can be
chosen orthogonal. There exists a unitary $U$ which transforms $A$ and $W$ into

$$U^*AU = \text{diag}(\lambda_1, \ldots, \lambda_r) \oplus A_1$$

and

$$U^*WU = \begin{pmatrix} 0 & 0 \\ 0 & W_1 \end{pmatrix}.$$  

$A_1$ has no eigenvalue $\mu$ with $f(\mu) = 0$. So by Theorem 2, $f_1(A_1) = W_1$ implies

$$\text{rank}(A^*_1|W_1) = n - r.$$  

As $S(A^*_1|W)$ and $S(A^{*_1}|W_1)$ are of equal rank, (5) follows.

The assumption on the elementary divisors of $A$ can not be dropped in general. Take $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $f(\lambda) = \bar{\lambda}\lambda$ as an example,

$$f_1(A) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = W.$$  

(5) is not satisfied, for there are two eigenvalues $\lambda$ of $A$ with $f(\lambda) = 0$, and

$$\text{rank}(A^*_1|W) = 1.$$  

3. SPECIAL CASES

There are two important special cases of Theorem 2 and 3.

**Theorem 4.** Let $A$ satisfy

(a) $A^*H + HA = W$

(b) $H - A^*HA = W$

where $H > 0$ and $W \geq 0$. Then

$$\text{Re}\lambda \geq 0, \quad |\lambda| < 1$$

for all eigenvalues $\lambda$ of $A$. If $A$ has $r > 0$ eigenvalues with

$$\text{Re}\lambda = 0, \quad |\lambda| = 1,$$

then the corresponding elementary divisors are linear and

$$n - r = \text{rank}(A^*_1|W).$$

Part (a) of Theorem 4 is also contained in [12, Corollary 4.1] or in [15]. It is known (see, e.g., [9] or [14]) that in (a) all eigenvalues of $A$ on the imaginary axis have linear elementary divisors. Using a lemma of Givens on
the field of values of \( A \), we can give a different proof of the linearity of those elementary divisors.

**Lemma 3.** (Givens [2]). *Let the field of values \( F(A) \) of \( A \) be defined as* \[ F(A) = \{ z | z = x^*Ax, x^*x = 1 \} \] *If \( \lambda \) is an eigenvalue of \( A \) lying on the boundary of \( F(A) \), then the elementary divisors corresponding to \( \lambda \) are linear.*

**Proof of Theorem 4.** Again there is no loss of generality if we work with
\[ A^* + A = W \succ 0, \]
\[ I - A^*A = W \succ 0. \]
If \( z = x^*Ax \), then (6) implies \( \text{Re} z = \frac{1}{2}(z + \overline{z}) = \frac{1}{2} x^*(A^* + A)x \succ 0 \). So an eigenvalue \( \lambda \) of \( A \) with \( \text{Re} \lambda = 0 \) is on the boundary of \( F(A) \) and Lemma 3 can be applied. Suppose now (7) holds. If \( z \in F(A) \), then \( |z|^2 = (x^*A^*x)(x^*Ax) \leq x^*A^*Ax \) for some \( x \) with \( x^*x = 1 \). By (7), \( x^*A^*Ax \leq 1 \), and any eigenvalue \( \lambda = e^{i\phi} \) of \( A \) is on the boundary of \( F(A) \).

Theorem 4b provides a new proof of the following result of Pták. Let \( \rho(A) \) be the spectral radius and \( \| A \| \) be the spectral norm of \( A \),
\[ \| A \| = \sqrt[\rho(A^*A)}. \]

**Theorem 5 [10].** *If \( A \) is an \( n \times n \) matrix with \( \| A \| = 1 \), then* \[ \| A^n \| = 1 \quad \Rightarrow \quad \rho(A) = 1. \]

**Proof.** From \( \| A \| = 1 \) one may deduce (7) and \( \rho(A) < 1 \). We have to show that there is an eigenvalue \( \lambda \) of \( A \) with \( |\lambda| = 1 \). Because of (7) and Lemma 2, the rank of \( S(A^* | V) \), \( V^2 = W \), will be studied.
\[
\text{rank} S(A^* | V) = \text{rank}(V, A^*V, \ldots, A^{n-1}V) \\
= \text{rank}(W + A^*WA + \ldots + A^{n-1}WA^{n-1}) \\
= \text{rank}[(I - A^*A) + (A^*A - A^{*2}A^2) + \ldots + (A^{n-1}A^{n-1} - A^{*n}A^n)] \\
= \text{rank}(I - A^{*n}A^n).
\]
\[ \| A^n \| = 1 \] implies that 1 is an eigenvalue of \( A^{*n}A^n \) or \( \text{rank}(I - A^{*n}A^n) < n \). Thus \( \text{rank} S(A^* | W) < n \), and by Theorem 4b at least one eigenvalue of \( A \) has modulus 1.
REMARK. The exponent $n$ in (8) is "critical", i.e., it cannot be replaced by a smaller one. Take the $n \times n$ matrix $B = (i,j+1)$ as an example [10]. We have $\|B\| = \|B^{n-1}\| = 1$ and $\rho(B) < 1$.

4. INEQUALITIES

In this section we give estimates for $f(\lambda)$, when $\lambda$ is an eigenvalue of $A$.

**Theorem 6.** Let $M$ be hermitian and $H > 0$, and let $d_1 \geq \cdots \geq d_n$ be the eigenvalues of $H^{-1}M$. Let $f$ be defined as in (4), and $\lambda$ be an arbitrary eigenvalue of $A$. If

$$
\sum_{i,j=0}^{n-1} c_{ij} A^i H A^j = M, \quad c_{ij} = c_{ji}
$$

holds, then

$$
d_n < f(\lambda) < d_1 \tag{9}
$$

**Proof.** Put

$$
\hat{A} = H^{1/2} A H^{-1/2} \quad \text{and} \quad \hat{M} = H^{-1/2} M H^{-1/2}, \tag{10}
$$

then

$$
\sum_{i,j=0}^{n-1} c_{ij} \hat{A}^i \hat{A}^j = \hat{M}.
$$

If $x$ is an eigenvector of $\hat{A}$ corresponding to $\lambda$, $\hat{A}x = \lambda x$, then

$$
x^* x \sum_{i,j=1}^{n-1} c_{ij} \hat{A}^i \lambda^j = x^* \hat{M} x
$$
or

$$
f(\lambda) = \frac{x^* \hat{M} x}{x^* x}. \tag{11}
$$

Since the eigenvalues of $\hat{M}$ are those of $H^{-1/2} \hat{M} H^{1/2} = H^{-1} M$, (9) follows from (11).

For the special case of the Lyapunov equation

$$
A^* H + H A = M, \quad H > 0, \tag{12}
$$
Kalman and Bertram [6] obtained
\[ d_n \leq 2 \text{Re} \lambda \leq d_1. \]
For (12) a set of sharper inequalities can easily be derived from a theorem of Fan [1].

**Definition** [3, p. 45]. Let \( (x) = (x_1, \ldots, x_n) \) and \( (y) = (y_1, \ldots, y_n) \) be two finite sequences of real numbers. We write
\[ (x_1, \ldots, x_n) < (y_1, \ldots, y_n), \]
if \( (x) \) and \( (y) \) can be arranged so as to satisfy the following three conditions:
\[
\begin{align*}
x_1 &> \cdots > x_n, \quad y_1 > \cdots > y_n \\
x_1 + \cdots + x_k &\leq y_1 + \cdots + y_k, \quad k = 1, 2, \ldots, n-1 \\
x_1 + \cdots + x_n &= y_1 + \cdots + y_n.
\end{align*}
\]

**Lemma 4** [1, 8]. If \( \lambda_1, \ldots, \lambda_n \) are the eigenvalues of \( A \) and \( \alpha_1, \ldots, \alpha_n \) are the eigenvalues of \( \frac{1}{2}(A^* + A) \), then
\[ (\text{Re} \lambda_1, \ldots, \text{Re} \lambda_n) < (\alpha_1, \ldots, \alpha_n). \]
(12) is equivalent to \( \hat{A}^* + \hat{A} = \hat{M} \) with \( \hat{A} \) and \( \hat{M} \) as in (10). From Lemma 4 we get the immediately the following relations.

**Theorem 7.** If (12) holds, and \( \lambda_1, \ldots, \lambda_n \) are the eigenvalues of \( A \) and \( d_1, \ldots, d_n \) the eigenvalues of \( H^{-1}M \), then
\[
\begin{align*}
(2 \text{Re} \lambda_1, \ldots, 2 \text{Re} \lambda_n) &< (d_1, \ldots, d_n) \\
(-2 \text{Re} \lambda_1, \ldots, -2 \text{Re} \lambda_1) &< (-d_1, \ldots, -d_1).
\end{align*}
\]

**References**


Received 6 November 1973