Some Remarks on the Spectra of Hermitian Matrices

J. Kovač-Striko
Faculty of Traffic
University of Zagreb
YU 41000 Zagreb, Yugoslavia

and

K. Veselić
Lehrgeb. Math. Physik
Fernuniversität Hagen
Postfach 940
D 5800 Hagen, West Germany

Submitted by Ludwig Elsner

ABSTRACT

The well-known Cauchy theorem connects the eigenvalues of a Hermitian matrix to the eigenvalues of a principal submatrix by a sequence of interlacing inequalities. In this note we derive some consequences of the assumption that some of these inequalities become equalities or near-equalities. Our results concern the tininess of the coupling off-diagonal block as well as corresponding subspace estimates.

Let A be an arbitrary $n \times n$ Hermitian matrix partitioned as

$$A = \begin{pmatrix} H & B^* \\ B & U \end{pmatrix}, \qquad H: m \times m. \tag{1}$$

Then according to the Cauchy interlacing theorem we have (cf. [3, §10.1])[1]

$$\alpha_1 \leqslant \theta_1, \dots, \qquad \alpha_m \leqslant \theta_m \tag{2}$$

---

[1] We shall use the notation of [3] throughout this paper.

LINEAR ALGEBRA AND ITS APPLICATIONS 145:221–229 (1991)

655 Avenue of the Americas, New York, NY 10010 0024-3795/91/$3.50

and

$$\alpha_n \geqslant \theta_m, \dots, \qquad \alpha_{n-m+1} \geqslant \theta_1 \tag{3}$$

where $\alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_n$ and $\theta_1 \leqslant \cdots \leqslant \theta_m$ are the eigenvalues of A and H, respectively.

Some applications gave rise to the following problem: What can be said about A if some of the inequalities in (2) or (3) become equalities or near-inequalities? The (plausible) answer is that then B has to be tiny. The results presented in this note are concerned with some aspects of this problem. The equality case is settled by the following.

THEOREM 1. *Let A and H be as above. Suppose that there is a set $\mathscr{S}$ of indices such that for every $k \in \mathscr{S}$*

$$\alpha_k = \theta_k \tag{4}$$

or

$$\alpha_{n-m+k} = \theta_k \tag{5}$$

holds. Then there is an orthonormal system u_k, $k \in \mathscr{S}$, such that

$$Bu_k = 0, \quad Hu_k = \theta_k u_k, \qquad k \in \mathscr{S}.$$

In particular, if (4) (*or* (5)) *holds for all θ_k, then $B = 0$.*

REMARK. We do not claim originality for this result (a fairly simple and elegant proof was supplied by the referee).[2] We include a proof of our own mainly because it gives some hints to treat the near-equality case below.

LEMMA 2. *Let*

$$A = \begin{pmatrix} 0 & B^* \\ B & U \end{pmatrix} \tag{6}$$

---

[2] This theorem can also be obtained as a corollary of a Hilbert-space result of Voigt and Weidmann [4] (some further results, related to the Cauchy theorem, are contained in [1, 2, 5]).

be a Hermitian matrix of order n, where the zero block is square of order p. If $\nu(A) \leqslant p-1$ or $\pi(A) \leqslant p-1$, then B has a nontrivial nullspace.

Proof of Lemma 2. If $p > n-p$, the lemma is trivial. For $p \leqslant n-p$ take any real $\eta \neq 0$; then $A-\eta$ is congruent to

$$\begin{pmatrix} -\eta & 0 \\ 0 & U-\eta+BB^*/\eta \end{pmatrix}.$$

If B had only a trivial nullspace, then rank $BB^* = p$ and $|\eta| \neq 0$ small enough would imply

$$\nu(A-\eta) \geqslant p.$$

This inequality would then necessarily extend to the case $\eta = 0$-a contradiction. This proves the lemma. (The second part of the assertion is obtained by considering $-A$ instead of A.) ■

Proof of Theorem 1. Without loss of generality we can assume that

$$H = \operatorname{diag}(\theta_1, \dots, \theta_m).$$

Then for $k \in \mathscr{S}$ we have $\nu(A-\alpha_k) \leqslant k-1$ and

$$A-\alpha_k = \begin{pmatrix} H_1-\alpha_k & 0 & 0 & B_1^* \\ 0 & H_2-\alpha_k & 0 & B_2^* \\ 0 & 0 & H_3-\alpha_k & B_3^* \\ B_1 & B_2 & B_3 & U-\alpha_k \end{pmatrix} \tag{7}$$

with

$$H_1-\alpha_k = \operatorname{diag}(\theta_1-\alpha_k, \dots, \theta_j-\alpha_k) \qquad \text{negative definite,}$$

$$H_2-\alpha_k = \operatorname{diag}(\theta_{j+1}-\alpha_k, \dots, \theta_k-\alpha_k) = 0,$$

$$H_3-\alpha_k = \operatorname{diag}(\theta_{k+1}-\alpha_k, \dots, \theta_m-\alpha_k).$$

Set

$$P = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ -B_1(H_1-\alpha_k)^{-1} & 0 & 0 & I \end{pmatrix}.$$

Then

$$P(A-\alpha_k)P^T = \begin{pmatrix} H_1-\alpha_k & 0 \\ 0 & \hat{A} \end{pmatrix},$$

$$\hat{A} = \begin{pmatrix} 0 & 0 & B_2^* \\ 0 & H_3-\alpha_k & B_3^* \\ B_2 & B_3 & Z \end{pmatrix},$$

$$Z = U - \alpha_k - B_1(H_1-\alpha_k)^{-1}B_1^*,$$

and by $\nu(\hat{A}) \leqslant k-j-1$ Lemma 2 applies. Thus, the matrix B_2 has a nontrivial nullspace. In other words, there is a unit vector u such that

$$Hu = \alpha_k u, \qquad Bu = 0,$$

which proves our theorem if card $\mathscr{S} = 1$. If card $\mathscr{S} > 1$ we proceed by deflation. Take a unitary matrix $\tilde{Q}$ of order $k-j$ such that

$$B_2' = B_2\tilde{Q}$$

has a vanishing last column. By setting

$$Q = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & \tilde{Q} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{pmatrix},$$

we have

$$Q^{-1}AQ = \begin{pmatrix} H & B'^* \\ B' & U \end{pmatrix},$$

where the kth column of B' vanishes. We now erase that row and column from the matrix above and repeat the whole procedure with another $k' \in \mathscr{S}$. This proves the theorem. ∎

The near-equality analog of Theorem 1 does not seem to be that easy to derive. To see where the difficulties lie, suppose that for some k

$$\theta_{k-1} < \alpha_k,$$

where $\theta_{-1} = -\infty$. In the notation of (7) (here $j = k-1$, $H_2 = \theta_k$, and $B_2 = b_2$ is a vector) the matrix $A - \alpha_k$ is congruent to

$$\begin{pmatrix} H_1 - \alpha_k & 0 & 0 & 0 \\ 0 & \theta_k - \alpha_k & 0 & b_2^* \\ 0 & 0 & H_3 - \alpha_k & B_3^* \\ 0 & b_2 & B_3 & Z_1 \end{pmatrix}, \tag{8}$$

where

$$Z_1 = U - \alpha_k - B_1(H_1 - \alpha_k)^{-1}B_1^*.$$

Now $\nu(H_1 - \alpha_k) = k-1$ and $\nu(A - \alpha_k) = k-1$ imply that the matrix

$$\begin{pmatrix} \theta_k - \alpha_k & 0 & b_2^* \\ 0 & H_3 - \alpha_k & B_3^* \\ b_2 & B_3 & Z_1 \end{pmatrix} \tag{9}$$

is positive semidefinite. Thus,

$$b_2 b_2^* \leqslant (\theta_k - \alpha_k)\left[U - \alpha_k + B_1(\alpha_k - H_1)^{-1}B_1^*\right]. \tag{10}$$

From this we can obtain the following estimate:

$$\|b_2\|^2 \leqslant (\theta_k - \alpha_k)\,\text{spread } A\left(1 + \frac{\text{spread } A}{4(\alpha_k - \theta_{k-1})}\right). \tag{11}$$

In fact, in (10) $\|U - \alpha_k\| \leqslant$ spread A, $\|B_1\| \leqslant \|A\|$. Since in the proof A can be replaced by $A - \alpha$, α real, we can use $\inf\|A - \alpha\| =$ spread $A/2$. Here the size of the spectral gap $\theta_k - \theta_{k-1}$ (note that $\alpha_k - \theta_{k-1} = \alpha_k - \theta_k + \theta_k - \theta_{k-1}$) plays a crucial role and spoils the estimate. The case $k = 1$ is an exception:

$$b_2 b_2^* \leqslant (\theta_1 - \alpha_1)(U - \alpha_1). \tag{12}$$

This leads to a simple and elegant estimate

$$\|b_2\|^2 \leqslant (\theta_1 - \alpha_1) \operatorname{spread} A. \tag{13}$$

(Note that here the matrix B_1 is void.) A further analysis of A by means of the Gaussian elimination and Sylvester inertia theorem gets more and more complicated, but it leads to the following.

CONJECTURE 3. Let A be as in (1), (2), (3), and let $\theta_{k-1} < \alpha_k$ for some k ($\theta_0 = -\infty$). Then for any $q \geqslant k$ there exists an orthonormal system $u_k, \dots, u_q$ such that

$$\sum_{i=k}^{q} \|Bu_i\|^2 \leqslant \operatorname{const} \sum_{i=k}^{q} (\theta_i - \alpha_i) \operatorname{spread} A, \tag{14}$$

where const depends only on n, m, k, q.

We shall now prove a spectral subspace estimate which will corroborate our conjecture.

THEOREM 4. *Let A, H be as in* (1), (2), (3), *and let* $\alpha_{k+1} > \alpha_k$ *for some k. Then the spectral subspaces* $\mathscr{A}_k$, $\mathscr{H}_k$ *spanned by the first k eigenvalues of A, H,*[3] *respectively, satisfy the inequality*

$$\sum_{i=1}^{k} \sin^2 \psi_i \leqslant \frac{1}{\alpha_{k+1} - \alpha_k} \sum_{i=1}^{k} (\theta_i - \alpha_i) \tag{15}$$

Here ψ_i *are the angles between* $\mathscr{A}_k$ *and* $\mathscr{H}_k$, *defined by* [3]

$$\cos \psi_j = \lambda_j\left[\left(F^*GG^*F\right)^{1/2}\right], \tag{16}$$

---

[3] Here and in the following we consider the spectral subspaces of A as naturally embedded in the whole space.

and F, G are the matrices containing orthonormal bases of $\mathscr{A}_k$, $\mathscr{H}_k$ *respectively.*

Proof. Again we suppose that $H = \operatorname{diag}(\theta_1, \dots, \theta_m)$ and write

$$A = \begin{pmatrix} H_1 & 0 & B_1'^* \\ 0 & H_2 & B_2'^* \\ B_1' & B_2' & U \end{pmatrix} = \begin{pmatrix} H_1 & B_1^* \\ B_1 & U_1 \end{pmatrix},$$

$H_1 = \operatorname{diag}(\theta_1, \dots, \theta_k)$, $B_1^* = (0, B_1'^*)$. Take a unitary Q which diagonalizes A:

$$Q^*AQ = \begin{pmatrix} A_1' & 0 \\ 0 & A_2' \end{pmatrix},$$

$$A_1' = \operatorname{diag}\left(\alpha_1, \dots, \alpha_k\right), \qquad A_2' = \operatorname{diag}\left(\alpha_{k+1}, \dots, \alpha_n\right).$$

Any such Q can be written as

$$Q = \begin{pmatrix} \sqrt{I - XX^*} & X \\ -X^* & \sqrt{I - X^*X} \end{pmatrix} \begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix},$$

where $\| X \| \leqslant 1$ and Q_1, Q_2 are some unitary blocks.

The matrices

$$G = \begin{pmatrix} I \\ 0 \end{pmatrix}, \qquad F = QG = \begin{pmatrix} \sqrt{I - XX^*} \\ -X^* \end{pmatrix} Q_1 \tag{17}$$

determine the spaces $\mathscr{H}_k, \mathscr{A}_k$ as linear spans of their rows, respectively. Obviously these spaces do not depend on Q_1. In particular, we have

$$F^*GG^*F = I - XX^*$$

such that

$$\sin^2 \psi_i = 1 - \lambda_i[I - XX^*] = \lambda_i[XX^*]$$

and

$$\sum_{i=1}^{k} \sin^2 \psi_i = \sum_{i=1}^{k} \lambda_i[XX^*] = \| X \|_F^2. \tag{18}$$

By setting

$$A_i = Q_i A_i' Q_i^*, \qquad i = 1,2,$$

we obtain

$$\begin{aligned} A &= \begin{pmatrix} H_1 & B_1^* \\ B_1 & U_1 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{I - XX^*} & -X \\ X^* & \sqrt{I - X^*X} \end{pmatrix} \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} \sqrt{I - XX^*} & X \\ -X^* & \sqrt{I - X^*X} \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{I - XX^*}\, A_1 \sqrt{I - XX^*} + XA_2X^* & \sqrt{I - XX^*}\, A_1 X - XA_2 \sqrt{I - X^*X} \\ X^*A_1 \sqrt{I - XX^*} - \sqrt{I - X^*X}\, A_2 X^* & X^*A_1X + \sqrt{I - X^*X}\, A_2 \sqrt{I - X^*X} \end{pmatrix}. \end{aligned} \tag{19}$$

Thus,

$$\begin{aligned} \sum_{i=1}^{k} \left(\theta_i - \alpha_i\right) &= \operatorname{Tr} H_1 - \operatorname{Tr} A_1 \\ &= \operatorname{Tr}\left(\sqrt{I - XX^*}\, A_1 \sqrt{I - XX^*} + XA_2X^*\right) - \operatorname{Tr} A_1 \\ &= \operatorname{Tr} X^*XA_2 - \operatorname{Tr} XX^*A_1. \end{aligned} \tag{20}$$

By (18) and

$$\operatorname{Tr} X^*XA_2 \geqslant \| X \|_F^2 \alpha_{k+1}, \qquad \operatorname{Tr} XX^*A_1 \leqslant \| X \|_F^2 \alpha_k,$$

we obtain

$$\sum_{i=1}^{k} \sin^2 \psi_i = \| X \|_F^2 \leqslant \frac{1}{\left(\alpha_{k+1} - \alpha_k\right)} \sum_{i=1}^{k} \left(\theta_i - \alpha_i\right) \tag{21}$$

This proves the theorem. ■

COROLLARY 5. *Under the conditions of Theorem 4 there is an orthonormal set $u_1, \dots, u_k$ such that*

$$Hu_i = \theta_i u_i, \qquad \sum_{i=1}^{k} \| Bu_i \|^2 \leqslant \frac{\sum_{i=1}^{k} \left(\theta_i - \alpha_i\right) \operatorname{spread} A}{\alpha_{k+1} - \alpha_k}. \tag{22}$$

Proof. From (19) we obtain

$$\sum_{i=1}^{k} \| Be_i \|^2 = \| B_1^* \|^2 \leqslant 2 \| X \|_F \max(\| A_1 \|, \| A_2 \|) \leqslant 2 \| X \|_F \| A \|.$$

Here, too, $\| A \|$ can be replaced by spread $A/2$, which proves the corollary. ∎

The estimate (22) shows that many spectral gaps are in fact not needed for estimating $\sum \| Bu_i \|^2$. (According to our conjecture the remaining gap $\alpha_{k+1} - \alpha_k$ is not needed either.)

For $k = 1$ (15) yields

$$\sin^2 \psi_1 \leqslant \frac{\theta_1 - \alpha_1}{\alpha_2 - \alpha_1}. \tag{23}$$

Here

$$\cos \psi_1 = | y_1^* z_1 | = \cos \varphi_1,$$

where φ_1 is the angle between the eigenvectors y_1, z_1 belonging to A, α_1 and H, θ_1, respectively. Now (23) coincides with the first of a series of estimates for the Ritz vectors obtained in [3, §11.9].

REFERENCES

1 J. Dancis and C. Davis, An interlacing theorem for eigenvalues of self-adjoint operators, *Linear Algebra Appl.* 88/89:117–122 (1987).
2 C. Davis, Beyond the minimax principle, *Proc. Amer. Math. Soc.* 81:401–405 (1981).
3 B. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J., 1980.
4 J. Voigt, Über die Friedrichsfortsetzung und die Monotonie der Eigenwerte, *Arch. Math. (Basel)* 26:183–188 (1975).
5 K. Fan and G. Pall, Imbedding conditions for Hermitian and normal matrices, *Canad. J. Math.* 9:298–304 (1957).

Received 12 December 1988; final manuscript accepted 29 May 1990