Eigenvalues of Perturbed Hermitian Matrices

C. C. PAIGE
McGill University
Montreal, Quebec, Canada

Communicated by Wallace Givens

ABSTRACT

It is a common experience that the perturbation, or even the omission, of some elements of a matrix often has negligible effect on some of the eigenvalues of the whole matrix. Here some new theorems are presented on this isolation effect in Hermitian matrices. The results are of importance in the computation of eigenvalues, particularly for tridiagonal matrices.

Let $A$, $F$, and $G$ be equivalently partitioned $n$ by $n$ Hermitian matrices

$$A = F + G, \quad F = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}, \quad G = \begin{pmatrix} 0 & A_{21}^H \\ A_{21} & 0 \end{pmatrix},$$

where $A_{11}$ is $k$ by $k$. Denote the eigenvalues of $A_{11}$ and $A_{22}$ by $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_k$ and $\mu_{k+1} \geq \mu_{k+2} \geq \cdots \geq \mu_n$ respectively. The eigenvalues $\mu_i(\theta)$ of $F + \theta G$ will be continuous functions of the scalar $\theta$, and so it is possible to order the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $A$, so that $\lambda_i = \mu_i(1), i = 1, 2, \ldots, n$. That is, the eigenvalues of $A$ can theoretically be ordered so that $\lambda_i$ "originates" from $\mu_i = \mu_i(\theta), \theta = 0$, as $\theta$ increases smoothly from 0 to 1.

If

$$||A_{21}||_2 = [\text{maximum eigenvalue of } (A_{21}^H A_{21})]^{1/2} = \epsilon,$$  \hspace{1cm} (2)

then the extreme eigenvalues of $G$ are $\pm \epsilon$, and from [1, p. 102] it follows, after some thought, that no eigenvalue $\mu_i(\theta)$ of $F + \theta G$ changes by more than $\pm \epsilon$ for $0 \leq \theta \leq 1$.  

Thus if the eigenvalues are wanted to within $\varepsilon_1$, only the two simpler
eigenvalue problems of $A_{11}$ and $A_{22}$ need be solved if $\varepsilon < \varepsilon_1$. Wilkinson
[1, p. 312] shows that a better result can often be obtained for a tridiagonal
Hermitian matrix; and this result will now be extended to any Hermitian
matrix. To do this the scalars $\gamma_i$ will be defined

$$\gamma_i \equiv \min_{j=k+1,\ldots,n} |\lambda_i - \mu_j|, \quad i = 1, 2, \ldots, k.$$  \hspace{1cm} (4)

Now by Schur’s theorem there exists a matrix $V$, unitary,

$$V = \begin{pmatrix} P & R \\ S & Q \end{pmatrix}, \quad V^H V = I, \hspace{1cm} (5)$$

(with the same partitioning as $A$), such that

$$V^H A V = D = \text{diag}(\lambda_i). \hspace{1cm} (6)$$

Then denoting the $i$th columns, $i \leq k$, of $V$, $P$, and $S$ by $v_i$, $p_i$, and $s_i$ respectively, it follows from Eqs. (5) and (6) that

$$Av_i = \lambda_i v_i, \quad v_i^H v_i = p_i^H p_i + s_i^H s_i = 1, \hspace{1cm} (7)$$

so that

$$(A_{22} - \lambda_i I)s_i = - A_{21} p_i, \hspace{1cm} (9)$$

giving

$$|s_i|_2 \leq \varepsilon|p_i|_2/\gamma_i, \quad \text{or} \quad |s_i|_2 \leq \varepsilon[\varepsilon^2 + \gamma_i^2]^{1/2} \leq 1, \hspace{1cm} (8)$$

ignoring the trivial case $\gamma_i = \varepsilon = 0$.

The bound (8) is important whenever $\varepsilon \ll \gamma_i$, as will now be shown. From Eqs. (7) and (1)

$$A_{11} p_i - \lambda_i p_i = - A_{21}^H s_i, \hspace{1cm} (9)$$

which combined with Eq. (8) and [1, p. 171] shows there is an eigenvalue
$\mu_j$ of $A_{11}$ such that

$$|\lambda_i - \mu_j| \leq \varepsilon^2/\gamma_i, \hspace{1cm} (10)$$

and this is superior to Eq. (3) whenever $\varepsilon < \gamma_i$. The bounds (3) and (10) then show that the error in taking any eigenvalue of $A_{11}$ as an approxima-
EIGENVALUES OF PERTURBED HERMITIAN MATRICES

Consideration to an eigenvalue of $A$ is bounded by $O(\varepsilon^2)$ unless the eigenvalue being approximated is also close to an eigenvalue of $A_{22}$, in which case the bound is only $\varepsilon$. Moreover it would seem that these are the best bounds obtainable in general, for consider the 2 by 2 matrix $A$ with $a_{11} = \alpha$, $a_{22} = \alpha + 2\delta$, and $a_{12} = a_{21} = \varepsilon$, then the eigenvalues of $A$ are

$$
\lambda = \alpha + \delta \pm (\delta^2 + \varepsilon^2)^{1/2},
$$

$$
= \alpha \pm \varepsilon \quad \text{if} \quad \delta = 0, \quad \text{cf. Eq. (3)},
$$

$$
= \alpha + (1 \pm \sqrt{2}) \delta \quad \text{if} \quad \varepsilon = \delta, \quad \text{cf. Eq. (10)}.
$$

To illustrate a possible practical use of Eq. (10) it is only necessary to quote the corresponding example used by Wilkinson [1, p. 312] whereby if the minimum separation of the eigenvalues of $A$ is known to be $10^{-2}$ then in a computation using 10 decimal places the submatrix $A_{22}$ can be ignored if $\varepsilon < 10^{-6}$ in (2).

This extension of Wilkinson's result might have been expected intuitively, and similarly it could be guessed that eigenvalues of $A$ originating from $A_{11}$ will be shielded from perturbations in $A_{22}$ by small $A_{21}$. Consider Eq. (1) and denote by $A + \delta A$ the matrix resulting from an Hermitian perturbation $\delta A_{22}$ in $A_{22}$ alone. Then from Eq. (6)

$$(A + \delta A)v_i - \lambda_i v_i = \delta A v_i,$$

so that there exists an eigenvalue $v_i$ of $A + \delta A$ satisfying

$$
|\lambda_i - v_i| \leq \|\delta A v_i\|_2 = \|\delta A_{22} s_i\|_2 \leq \|\delta A_{22}\|_2.
$$

But (8) can be used to give the better bound

$$
|\lambda_i - v_i| \leq \varepsilon \|\delta A_{22}\|_2 (\varepsilon^2 + \gamma_i^2)^{1/2} \leq \varepsilon \|\delta A_{22}\|_2 / \gamma_i.
$$

This last result, superior though it is to the well known result (11), may in many cases be improved even further. Here the Rayleigh quotient $\rho_i$ of $A + \delta A$ and $v_i$ is seen to be

$$
\rho_i = \lambda_i + v_i^H \delta A v_i = \lambda_i + s_i^H \delta A_{22} s_i,
$$

so that

$$
y_i \equiv (A + \delta A) v_i - \rho_i v_i = \delta A v_i - (s_i^H \delta A_{22} s_i) v_i,
$$
gives
Now if there is only one eigenvalue $v_j$ of $A + \delta A$ such that

$$|v_j - \rho_i| < \zeta_i$$

then following [1, p. 188]

$$|v_j - \rho_i| \leq y_i^H y_i / \zeta_i,$$

i.e.

$$|v_j - \lambda_i - s_i^H \delta A_{22}s_i| \leq \|\delta A_{22}s_i\|_2 / \zeta_i,$$

or

$$|\lambda_i - v_j| \leq \frac{\epsilon^2 (1 + \|\delta A_{22}\|_2 / \zeta_i)}{(\epsilon^2 + \gamma_i^2)}$$

from Eq. (8). This result is important as it shows that for some well separated eigenvalues of $A$ the effect of an Hermitian perturbation in $A_{22}$ will be diminished by $O(\|A_{22}\|_2^2)$. This is a very satisfying result, as among other things, Eqs. (12) and (15) explain why small eigenvalues of very large Hermitian matrices are often surprisingly well-conditioned. As an example consider the matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \frac{10^{-5}}{10^{-5}} \frac{10^{-10}}{10^{-10}},$$

which by Eq. (3) has eigenvalues 3, 1, $10^{-10}$, to within $10^{-5}$, and using Eq. (10) it can be seen that these are accurate to $O(10^{-10})$. Now on a floating point arithmetic computer where a number $x$ is stored as

$$\bar{x} = x(1 + \eta), \quad |\eta| \leq 10^{-10},$$

the best eigenvalue algorithms will only introduce relative errors of $O(10^{-10})$ in each element. Thus consider a perturbation of $O(10^{-10})$ in the leading 2 by 2 submatrix. The usual result (11) suggests that the resulting smallest eigenvalue could be meaningless, but the new result (12) shows that it will not be in error by more than $O(10^{-15})$. However any reasonable algorithm will take account of symmetry, and for a symmetric perturba-
tion Eq. (15) shows that the error in the smallest eigenvalue will be no more than $O(10^{-20})!$ This is a type of computational result that is often observed in practice.

This isolating effect caused by small off-diagonal submatrices turns out to be cumulative in a sense that can now be described. The analysis is quite simple and need only be given for a 3 by 3 partitioning of $A$ and $V$, thus, for $i \leq k$, let

$$A = \begin{pmatrix} A_{11} & A_{12}^H & A_{13}^H \\ A_{21} & A_{22} & A_{23}^H \\ A_{31} & A_{32} & A_{33} \end{pmatrix}, \quad v_i = \begin{pmatrix} p_i \\ s_i \\ t_i \end{pmatrix}, \quad B = \begin{pmatrix} A_{11} & A_{12}^H \\ A_{21} & A_{22} \\ A_{31} & A_{32}^H \end{pmatrix},$$

$$C = \begin{pmatrix} A_{22} & A_{23}^H \\ A_{32} & A_{33} \end{pmatrix},$$

(16)

where again

$$Av_i = \lambda_i v_i, \quad v_i^H v_i = 1,$$

(17)

and the eigenvalues $\lambda_i$ of Hermitian $A$ are related to the diagonal blocks as before. Let

$$\varepsilon_{21} = ||A_{21}||_2, \quad \varepsilon_{31} = ||A_{31}||_2, \quad \varepsilon_{32} = ||A_{32}||_2,$$

and let $\beta_i, \gamma_i,$ and $\delta_i$ denote the minima of the moduli of the eigenvalues of $B - \lambda_i I$, $C - \lambda_i I$, and $A_{33} - \lambda_i I$ respectively [cf. Eq. (4)]. Then from the last $m$ rows of Eq. (17), using Eq. (16),

$$(A_{33} - \lambda_i I)t_i = -A_{31}\beta_i - A_{32}s_i.$$

while from the last $l + m$ rows of Eq. (17)

$$(C - \lambda_i I)\begin{pmatrix} s_i \\ t_i \end{pmatrix} = -\begin{pmatrix} A_{21} \\ A_{31} \end{pmatrix} p_i.$$

so that

$$||t_i||_2 \leq \left(\varepsilon_{31} + \varepsilon_{32}^2 ||p_i||_2^2/\gamma_i^2\right).$$

(18)
This bound is the extended equivalent of Eq. (8), and can now be used to bound the effect that $A_{33}$, or any perturbation in $A_{33}$, has on the eigenvalue $\lambda_i$ of $A$ originating from $A_{11}$. Thus from the first $k + l$ rows of Eq. (17), with Eq. (16)

$$(B - \lambda_i I) \begin{pmatrix} p_i \\ s_i \end{pmatrix} = - \begin{pmatrix} A_{31}^H \\ A_{32}^H \end{pmatrix} t_i = y_i, \text{ say,}$$

so that

$$||y_i||_2^2 \leq (\varepsilon_{31}^2 + \varepsilon_{32}^2)||t_i||_2^2,$$

which combines with (18) to show that $\beta_i$, the modulus of the error in taking an eigenvalue of $B$ to approximate this eigenvalue $\lambda_i$ of $A$, satisfies

$$\beta_i \leq (\varepsilon_{31}^2 + \varepsilon_{32}^2)^{1/2}(\varepsilon_{31} + \varepsilon_{32}(\varepsilon_{21}^2 + \varepsilon_{31}^2)^{1/2}/\gamma_i)/\delta_i,$$

$$\leq \varepsilon_{32}^2 \varepsilon_{21}/(\gamma_i \delta_i), \text{ if } \varepsilon_{31} = 0. \quad (19)$$

It is clear from Eq. (19) that $A_{31}$ in Eq. (16) must be significantly smaller than $A_{21}$ or $A_{32}$ for the isolation effect to be magnified. For matrices of tridiagonal and other narrow band form, $A_{31}$ can be chosen to be zero, and here Eq. (19) has most significance. Thus for a tridiagonal matrix several fairly small next to diagonal elements have a multiplicative effect that isolates some eigenvalues from distant matrix elements, as a result several eigenvalues can often be found to almost machine accuracy by considering a truncated portion of the matrix only, even when there are no very small next to diagonal elements. This is particularly noticeable in the Lanczos, Givens, and Householder tridiagonalizations of Hermitian $A$, where it is known that the extreme eigenvalues of $T_k$, the leading $k$ by $k$ portion of the tridiagonal matrix, converge very swiftly with increasing $k$ to the corresponding extreme eigenvalues of $A$.

However it is not true that if $\gamma_i$ and $\delta_i$ are very small in Eq. (19) then an effect of $O(\varepsilon^2)$ can be obtained as might be hoped, for consider the 3 by 3 matrix with diagonal elements $\alpha$ and next to diagonal elements $\varepsilon$; this has eigenvalues $\alpha, \alpha \pm \sqrt{2}\varepsilon$, while each 2 by 2 diagonal submatrix has eigenvalues $\alpha \pm \varepsilon$. Thus not all eigenvalues need be equally isolated from "distant" effects.

The results given so far can be summarized and extended for the most interesting particular case of block tridiagonal matrices, and this will now be done using the following notation. Let $A$ be an $n \times n$ Hermitian
block tridiagonal matrix with eigenvector $v$ partitioned conformably as

$$A = \begin{bmatrix}
A_1 & E_2^H \\
E_2 & A_2 & E_3^H \\
& \ddots & \ddots \\
E_{r-1} & A_{r-1} & E_r^H \\
& E_r & A_r
\end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\
v_2 \\
\vdots \\
v_{r-1} \\
v_r \end{bmatrix}, \quad (20)$$

with

$$Av = \lambda v, \quad v^H v = 1, \quad (21)$$

where $A_i$ is $n_i \times n_i$, $i = 1, 2, \ldots, r$, and $\epsilon_i \equiv ||E_i||_2$, $i = 2, 3, \ldots, r$. For $i = 1, 2, \ldots, r$, let $m_i \equiv n_i + n_{i+1} + \cdots + n_r$, and let $u_i$ be the vector of the last $m_i$ elements of $v_i$, and $T_i$ be the last $m_i \times m_i$ principal submatrix of $A$, and define $\tau_i$ to be the least distance from $\lambda$ in Eq. (21) to an eigenvalue of $T_i$. Then from the last $m_i$ rows of Eq. (21) for $i = 2, 3, \ldots, r$

$$(T_i - \lambda I)u_i = - \begin{bmatrix} E_i v_{i-1} \\
0 \end{bmatrix},$$

so that

$$||u_i||_2 \leq ||u_{i-1}||_2 \leq ||u_{i-1}||_2 \epsilon_i / \tau_i,$$

giving

$$||v_r||_2 \leq ||v_1||_2 \epsilon_2 \epsilon_3 \cdots \epsilon_r / (\tau_2 \tau_3 \cdots \tau_r). \quad (22)$$

This result will be used to show the extent to which eigenvalues originating from $A_1$, along with their eigenvectors, are shielded from changes in $A_r$ and $E_r$. Equation (22) holds for any eigenvalue $\lambda$ of $A$ as long as $\tau_2 \tau_3 \cdots \tau_r$ is nonzero, but the bound can only be expected to be small for some of the eigenvalues originating from $A_1$. Note that a more complicated bound could also be found in terms of $\epsilon_i$ and $\delta_i$, $i = 2, 3, \ldots, r$, where $\delta_i$ is the least distance from $\lambda$ to an eigenvalue of $A_i$.

We first examine the effect of approximating $\lambda$ and $v$ by a corresponding pair from a submatrix of $A$.

**Theorem.** Let $B$ be the $m \times m$ matrix obtained by deleting the last $n_r$ rows and columns of $A$, and $w$ the vector obtained by deleting the last $n_r$ elements of $v$. Suppose
\[ Bz_i = \mu_i z_i, \quad z_i^H z_k = \delta_{ik}, \quad i, k = 1, 2, \ldots, m, \quad (23) \]

and
\[ w = \sum_{i=1}^{m} \alpha_i z_i, \quad (24) \]

Then with the notation in Eqs. (20) and (21)

(i) \[ |\mu_j - \lambda| \equiv \min_i |\mu_i - \lambda| \leq \varepsilon_r \|v_r\|_2/(1 - \|v_r\|_2^2)^{1/2}, \quad (25) \]

while if
\[ \delta \equiv \min_{i \neq j} |\mu_i - \lambda|, \quad (26) \]

(ii) \[ \left\| v - \alpha_j \begin{bmatrix} x_j \\ 0 \end{bmatrix} \right\|_2^2 \leq (1 + \varepsilon_r^2/\delta^2) \|v_r\|_2^2, \quad (27) \]

and finally if \[ |\mu_i - w^H B w| \geq \zeta \text{ for } i \neq j, \]

(iii) \[ |\mu_j - \lambda| \leq (\|A_r - \lambda I\|_2 + \varepsilon_r^2/\zeta) \|v_r\|_2^2/(1 - \|v_r\|_2^2). \quad (28) \]

Proof. From the leading \( n - n_r \) elements of Eqs. (20) and (21)
\[ f \equiv (B - \lambda I)w = - \begin{pmatrix} 0 \\ E_r v_r \end{pmatrix}, \quad w^H w = 1 - v_r^H v_r, \quad (29) \]

so that
\[ \|f\|_2^2 = \sum_s |\alpha_s|^2 |\mu_i - \lambda|^2 \geq \delta^2 \sum_{s \neq j} |\alpha_s|^2. \]
\[ \therefore \left\| w - \alpha_j z_j \right\|_2^2 = \sum_{s \neq j} |\alpha_s|^2 \leq \varepsilon_r^2 \|v_r\|_2^2 / \delta^2 \]

from which Eq. (27) is seen to hold. Finally defining the Rayleigh quotient \( \rho \) and using Eq. (29)
\[ \rho \equiv \frac{w^H B w}{w^H w} = \lambda - \frac{v_r^H E_r v_{r-1}}{w^H w}; \quad (30) \]
but from the last $n_r$ rows of Eq. (21)

$$E_r v_{r-1} = -(A_r - \lambda I) v_r.$$  

$$\therefore \quad |\rho - \lambda| \leq ||A_r - \lambda I||_2 ||v_r||_2^2 / ||v||_2^2.$$  

Now from Eqs. (29) and (30) we certainly have that

$$||Bw - \rho w||_2 \leq \varepsilon_r ||v_r||_2,$$

since the Rayleigh quotient minimizes the norm on the left-hand side. Thus from [1, p. 188] it follows that

$$|\mu_j - \rho| \leq \frac{\varepsilon_r^2 ||v_r||_2^2}{\zeta (1 - ||v_r||^2)},$$

and since $|\mu_j - \lambda| \leq |\mu_j - \rho| + |\rho - \lambda|$, the result (28) follows. 

Inserting the bound (22) in these results shows the isolating effect caused by small off-diagonal blocks. Equations (25) and (28) give different order bounds on the error in taking the eigenvalue $\mu_j$ of $B$ in place of the eigenvalue $\lambda$ of $A$, while Eq. (27) shows how well $[z_j^T, 0]$ approximates $v^T$. As is usual the eigenvector bound is not as good as the eigenvalue bound. The practical computational problem of deciding when a symmetric tridiagonal matrix can be considered as two matrices of smaller dimensions for the purpose of computing eigenvalues, but not eigenvectors, has been discussed for example by Kahan [2].

We now consider the effect of a perturbation in $A_r$ on $\lambda$.

**Theorem.** Let $A$ and $v$ be as given in Eqs. (20) and (21), and consider the addition of an Hermitian perturbation $\delta A_r$ to $A_r$, then the resulting matrix $A + \delta A$ has an eigenvalue $v$ such that

$$|\lambda - v| \leq ||\delta A_r v_r||_2 \leq ||\delta A_r||_2 ||v_r||_2.$$

(31)

If $v$ is the only eigenvalue of $A + \delta A$ within a distance $\zeta$ of $\rho$ where

$$\rho \equiv v^H (A + \delta A) v,$$

then

$$|\lambda - v| \leq ||\delta A_r||_2 (1 + ||\delta A_r||_2 / \zeta)||v_r||^2.$$

(33)
Proof. From Eq. (21)

\[(A + \delta A)v - \lambda v = \delta Av, \quad ||\delta Av||_2 = ||\delta A_r v_r||_2.\]  

(34)

and \(A + \delta A\) is Hermitian, so Eq. (31) follows directly from [1, p. 171]. Now in Eq. (32)

\[\rho = \lambda + v^H \delta A v = \lambda + v_r^H \delta A_r v_r,\]  

(35)

and from Eq. (34) and the optimality of the Rayleigh quotient

\[||(A + \delta A)v - \rho v||_2 \leq ||\delta A_r v_r||_2,\]  

so following [1, p. 188]

\[|v - \rho| \leq ||\delta A_r v_r||_2 \sigma_0 \zeta,\]  

(36)

and Eq. (33) follows using \(|v - \lambda| < |v - \rho| + |\rho - \lambda|\) with Eqs. (35) and (36). A weaker bound could also be found for the change in the eigenvector.

Once more the bound (22) can be used with (31) and (33) to show how eigenvalues arising in one part of the matrix can be sheltered from perturbations in another part by small off-diagonal blocks. This isolating effect is seen to be strong for some eigenvalues, and this explains why the smaller eigenvalues of some tridiagonal matrices with some very large elements can be found remarkably accurately. The reader only has to consider the 10 by 10 tridiagonal matrix with diagonal elements \((11 - i)^{10}, i = 1, 2, \ldots, 10\), and next to diagonal elements of unity, an example suggested by J. H. Wilkinson.

REFERENCES


Received March, 1970; revised December, 1972