Matrix perturbations: bounding and computing eigenvalues
Reis da Silva, R. J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 2

Hermitian matrices

When $A$ is Hermitian and $E$ is Hermitian positive semi-definite, Weyl’s inequalities (Theorem 1.1.7) show that the eigenvalues of $A$ are lower bounds for the eigenvalues of $A + E$. For a long time these were the best available bounds. More recently, however, a result by Ipsen and Nadler [46] improves Weyl’s inequalities. Both (Weyl’s and the result from [46]) offer cheaply computable lower bounds for the eigenvalues of perturbed Hermitian matrices.

In this chapter we improve the bounds from Ipsen and Nadler [46] for the eigenvalues of $A + E$. The starting point is a Hermitian, positive semi-definite, rank-one perturbation, $E$, of a Hermitian matrix, $A$. The bounds proposed rely on the construction of a sequence of matrices approximating $A$ in a certain sense. Each element of this sequence determines an element of a sequence of lower bounds for each eigenvalue of $A + E$. The number of available eigenvalues and eigenvectors of the original matrix determines how close those approximations can be. If the perturbation is of low rank, the new bounds are relatively inexpensive to compute. For the smallest eigenvalue, the first two elements of the sequence correspond to Weyl’s and Ipsen and Nadler’s bounds. For the other eigenvalues, however, our approach is different from that of [46]. We then use the same procedures to bound eigenvalues of bordered diagonal matrices as well as singular values of rank-$k$ perturbations and other modifications of $n \times m$ matrices.
2.1 Introduction

The need to bound the eigenvalues of perturbed matrices arises, for instance, in problems related to the discretization of differential equations. Consider the following reaction-diffusion equation

\[- u'' + su = \lambda u \quad \text{in } I = [0, 1], \quad \text{with } u(0) = u(1) = 0. \quad (2.1)\]

In Equation (2.1), let the function \( s : [0, 1] \rightarrow \mathbb{R} \) be a continuous nonnegative local perturbation, which is zero everywhere except in the interval \((\frac{1}{2} - \xi, \frac{1}{2} + \xi)\) for some \( \xi \in [0, \frac{1}{2}] \) where it is defined by

\[ s(x) = e^{-g(x)} > 0 \quad \text{with} \quad g(x) = \frac{\xi^2}{(x - (1/2 - \xi))(x - (1/2 + \xi))}. \]

Take \( n \in \mathbb{N} \) and define the mesh width \( h = 1/n \) and the grid points \( x_q = qh \) for \( q \in \{0, \ldots, n\} \). The Finite Difference discretization of Problem (2.1) on the grid points \( x_0, \ldots, x_n \) yields a positive definite Hermitian tridiagonal matrix \( A \), resulting from the diffusion part and a Hermitian diagonal matrix \( E \), resulting from \( s \), whose rank is dependent on \( h \) and the variable \( \xi \).

The problem

In this chapter we are interested in lower bounds for the eigenvalues of discrete problems like those of Equation (2.1), i.e., lower bounds for the eigenvalues of \( A + E \). We may state it as follows.

Let \( A, E \in \mathbb{C}^{n \times n} \) be Hermitian matrices with the additional property that \( E \) is of rank-\( k \). Assuming that the eigendata of the matrices \( A \) and \( E \) are known, determine (lower) bounds for the eigenvalues of \( A + E \).

Because, at least in theory, we can always apply a unitary similarity transformation that diagonalizes \( A \) to both \( A \) and \( E \) we may assume, without loss of generality, that \( A \) is diagonal. The general case will be discussed later, in Section 2.2.3.
For \( q \in \{1, \ldots, n\} \) consider the \( n \times n \) matrix

\[
X_q = \begin{bmatrix}
\Lambda_q \\
\lambda_q I_{n-q}
\end{bmatrix}
\]

where \( \Lambda_q = \begin{bmatrix}
\lambda_1 \\
\cdot & \cdot \\
\lambda_q
\end{bmatrix} \), \hspace{1cm} (2.2)

and \( \lambda_q \) stands for the \( q \)th eigenvalue (in non-decreasing order) of \( A \). Observe that when \( q = 1 \) and \( q = n \) we have \( X_1 = \lambda_1 I_n \) and \( X_n = A \), respectively. Moreover,

\[
x^*(X_p - X_q)x \geq 0, \quad \text{for all } x \in \mathbb{C}^n, \quad \text{and } \quad p, q \in \{1, \ldots, n\} \quad \text{with} \quad p \geq q
\]

and we say that \( X_q \) is an approximation from below to \( X_p \). In particular, each \( X_q \) approximates \( A \) from below. Consequently, it follows from Weyl’s theorem (Theorem 1.1.7) that for \( j \in \{1, \ldots, n\} \)

\[
\lambda_j(X_1 + E) \leq \lambda_j(X_2 + E) \leq \ldots \leq \lambda_j(X_{n-1} + E) \leq \lambda_j(A + E).
\]

\hspace{1cm} (2.3)

**Remark 2.1.1.** We can obtain upper bounds for the eigenvalues by using approximations \( X_q \) from above, i.e, such that \( X_q - A \) is positive semi-definite, or by applying the theory that follows to \( -A \) instead of \( A \).

The following example illustrates, for two small matrices \( A \) and \( E \), what we have just discussed.

### 2.1.1 Example

Let \( w := [1, 1, 1, 1, 1]^* \) and consider the symmetric rank-one matrix \( E := ww^* \) and the real diagonal matrix

\[
A = \begin{bmatrix}
2 & 3 \\
3 & 3.1 \\
& & 4 \\
& & & 4.1
\end{bmatrix} = X_5
\]
as well as

\[
X_1 = \begin{bmatrix}
2 & 2 & 2 & 2 & 2 \\
2 & 2 & 2 & 2 & 2 \\
2 & 2 & 2 & 2 & 2 \\
\end{bmatrix},
X_2 = \begin{bmatrix}
2 & 3 & 3 & 3 & 3 \\
3 & 3 & 3 & 3 & 3 \\
\end{bmatrix},
X_3 = \begin{bmatrix}
2 & 3 & 3.1 & 3.1 & 3.1 \\
3 & 3.1 & 3.1 & 3.1 & 3.1 \\
\end{bmatrix},
X_4 = \begin{bmatrix}
2 & 3 & 3.1 & 3.1 & 3.1 \\
3 & 3.1 & 3.1 & 3.1 & 3.1 \\
\end{bmatrix}.
\]

The eigenvalues of \(X_q + E\) for each \(q \in \{1, \ldots, 5\}\) are displayed in Table 2.1.

<table>
<thead>
<tr>
<th></th>
<th>(\lambda_1)</th>
<th>(\lambda_2)</th>
<th>(\lambda_3)</th>
<th>(\lambda_4)</th>
<th>(\lambda_5)</th>
<th>(|X_q - X_{q-1}|)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_1 + E)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>(X_2 + E)</td>
<td>2.1716</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>7.8284</td>
<td>1</td>
</tr>
<tr>
<td>(X_3 + E)</td>
<td>2.1821</td>
<td>3.0250</td>
<td>3.1</td>
<td>3.1</td>
<td>7.8929</td>
<td>0.1</td>
</tr>
<tr>
<td>(X_4 + E)</td>
<td>2.2201</td>
<td>3.0473</td>
<td>3.5089</td>
<td>4</td>
<td>8.3236</td>
<td>0.9</td>
</tr>
<tr>
<td>(A + E)</td>
<td>2.2214</td>
<td>3.0475</td>
<td>3.5291</td>
<td>4.0519</td>
<td>8.3503</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 2.1: Eigenvalues of \(X_q + E\) for \(q \in \{1, \ldots, 5\}\)

Observe that each \(X_q\) is defined using the \(q\) smallest eigenvalues of \(A\) and that as \(q\) increases, all eigenvalues of \(X_q + E\) move, indeed, closer to those of \(A + E\) as showed by (2.3). Unfortunately, the effort involved in the computation of the eigenvalues of \(X_q + E\) seems the same as that involved in the computation of those of \(A + E\). If that obstacle could be removed, however, then the eigenvalues of \(X_q + E\) could be used also in practice as lower bounds for the eigenvalues of \(A + E\).

\textbf{Remark 2.1.2.} \textit{In the rank-one situation and for given \(j \in \{1, \ldots, n - 1\}\), only if \(q > j\) is the bound for \(\lambda_j\) due to \(X_q\) an improvement over Weyl’s bound.}
2.2 Hermitian rank-$k$ perturbations

Example 2.1.1 shows that increasingly accurate lower bounds for the eigenvalues of $A + E$ are easily defined. The error matrix resulting from approximating $A + E$ by $X_q + E$, is

$$S_q = A + E - (X_q + E) = \begin{bmatrix} \Lambda_q \\ \Lambda_{n-q} \end{bmatrix} - \begin{bmatrix} \Lambda_q \\ \lambda_q I_{n-q} \end{bmatrix} = \begin{bmatrix} O_q \\ L_q \end{bmatrix}$$

where $L_q = \Lambda_{n-q} - \lambda_q I_{n-q}$. More specifically, when considering two succeeding matrices $X_q$ and $X_{q+1}$ we obtain

$$\|S_{q+1}\|_2 = \|S_q\|_2 - (\lambda_{q+1} - \lambda_q). \tag{2.4}$$

Equation (2.4) estimates the difference expected from using each next element of the sequence. In other words, if $A$ has a multiple eigenvalue, $\lambda_q$, of multiplicity $m$, then all the matrices between $X_q$ and $X_{q+m}$ are equal and, thus, have the same eigenvalues. In this case, the bounds do not change. On the other hand, if the gap between two consecutive eigenvalues is large, we may hope to do better by using the next matrix.

Although the eigenvalues of $X_q + E$ provide lower bounds for the eigenvalues of $A + E$, the reader might be inquiring the usefulness of Example 2.1.1. For any $q \in \{1, \ldots, n\}$, the dimension of the matrix $X_q + E$ is the same as that of $A + E$ and the computation of the eigenvalues of $X_q + E$ seem as costly as the computation of the eigenvalues of $A + E$. Fortunately, as Proposition 2.2.1 will show, that need not be the case.

Proposition 2.2.1 focuses on the perturbation of multiple eigenvalues and, in short, states that a Hermitian rank-one perturbation of a Hermitian matrix, $X$ affects at most single copies of multiple eigenvalues. The remaining copies are not perturbed. The reasoning is that if $X$ has an eigenvalue $\lambda$ of multiplicity $m$ we can find $m - 1$ eigenvectors associated with $\lambda$ orthogonal to the perturbation (eigen)vector $w$. Proposition 2.2.1 suggests, thus, that the computation effort should be directed to those eigenvalues of $X_q + E$ that are not eigenvalues of $X_q$. 
Proposition 2.2.1. Let $\sigma \in \mathbb{R}$, $w \in \mathbb{C}^n$ with $\|w\|_2 = 1$ and define $E := w\sigma w^*$. Let $X \in \mathbb{R}^{n \times n}$ be diagonal with exactly $t$ distinct eigenvalues denoted by $\lambda_p$ ($1 \leq p \leq t$). Write $P_p$ for the orthogonal projection onto the eigenspace $X_p$ corresponding to $\lambda_p$ and define $\theta_p$ by

$$\|P_p w\|_2 = \cos \theta_p \quad \text{where} \quad \theta_p = \angle(w, X_p) \quad \text{is chosen in} \quad \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Define yet the $t \times t$ matrix

$$Y = \begin{bmatrix} \lambda_1 & \cdots & \vdots \\ \vdots & \ddots & \vdots \\ \lambda_t & \cdots & \lambda_t \end{bmatrix} + y\sigma y^*, \quad \text{where} \quad y = \begin{bmatrix} \cos \theta_1 \\ \vdots \\ \cos \theta_t \end{bmatrix}.$$

Then the eigenvalues of $X + E$ are the eigenvalues of $X$, of which single copies of multiple eigenvalues are replaced by the eigenvalues of $Y$.

Proof. We give the details of the proof for $t = 2$ after which the proof for $t > 2$ should be clear. Let

$$X = \begin{bmatrix} \lambda_1 I_{m_1} \\ \lambda_2 I_{m_2} \end{bmatrix}$$

where $m_1 > 1$ is the multiplicity of $\lambda_1$, $m_2 > 1$ is the multiplicity of $\lambda_2$ and $m_1 + m_2 = n$. For unitary matrices $Q_1 \in \mathbb{C}^{m_1 \times m_1}$ and $Q_2 \in \mathbb{C}^{m_2 \times m_2}$ define

$$Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \quad \text{and thus}, \quad Q^*XQ = X.$$

Now, since each $P_p w \in X_p$ we can choose each $Q_p$ such that $Q^* w$ has at most two nonzero entries:

$$Q^* w = \cos \theta_1 e_1 + \cos \theta_2 e_{m_1 + 1}.$$

Thus,

$$Q^*(X + E)Q = X + Q^*w\sigma w^*Q.$$

Note that the second term of the right-hand side has nonzero entries only at two rows and corresponding two columns. By appropriate permutations, those columns


and rows can be placed at the top left. Both $\lambda_1$ and $\lambda_2$ continue to be eigenvalues of $X + E$ now with multiplicity $m_1 - 1$ and $m_2 - 1$ respectively. As for the remaining two eigenvalues they have been replaced by those of $Y$:

$$Y = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} + \begin{bmatrix} \cos \theta_1 \\ \cos \theta_2 \end{bmatrix} \sigma [\cos \theta_1, \cos \theta_2].$$

This proves the statement for $t = 2$. For $t > 2$ the proof is similar.

If the rank of $E$ is larger than one, a similar situation occurs. This is expressed in Corollary 2.2.2.

**Corollary 2.2.2.** If, in the conditions of Proposition 2.2.1, the matrix $E$ is Hermitian and of rank-$k$ (with $1 \leq k \leq n$), then at most $k$ copies of each distinct eigenvalue of $X$ are not eigenvalues of $X + E$. Therefore, at most $kt$ of the total number of eigenvalues of $X$ are perturbed.

We conclude from Example 2.1.1 and from the two previous results that only the eigenvalues of $X_q + E$ that are not eigenvalues of $X_q$ need to be computed. It remains unclear, however, how that can be achieved. We address this problem in the next section.

### 2.2.1 Computing the bounds

Let $W \in \mathbb{C}^{n \times k}$ and $\Sigma \in \mathbb{R}^{k \times k}$ be such that

$$E = W \Sigma W^*$$

is a rank-$k$ Hermitian matrix. Consider, now, the following procedure.
Procedure 2.2.3. For any $j \in \{0, \ldots, n-k\}$ partition the $n \times k$ matrix $W$ as $W_j$, consisting of the first $j$ rows and $W_{n-j}$, consisting of the last $n-j$ rows, i.e.,

$$W = \begin{bmatrix} W_j \\ W_{n-j} \end{bmatrix}.$$ 

Therefore, if $Q \in \mathbb{C}^{(n-j)\times(n-j)}$ and $R \in \mathbb{C}^{k\times k}$ are such that

$$Q \begin{bmatrix} R \\ O \end{bmatrix} = W_{n-j} = W_j$$

is a QR-decomposition of $W_{n-j}$, we have for the unitary $G = I_j \oplus Q$ that

$$G^*EG = \begin{bmatrix} E_{j+k} \\ O \end{bmatrix}$$

where $E_{j+k} = \begin{bmatrix} W_j \\ R \end{bmatrix} \Sigma \begin{bmatrix} W_j^* \\ R^* \end{bmatrix}$

is a $(j+k) \times (j+k)$ matrix.

The goal of this procedure is to introduce a large zero trailing principal submatrix using a unitary transformation $G$ that leaves $\Lambda_q$ invariant, i.e, such that $G^*\Lambda_qG = \Lambda_q$. We can now use Procedure 2.2.3 to prove the following result.

Theorem 2.2.4. Let $E \in \mathbb{C}^{n\times n}$ be the Hermitian rank-$k$ matrix defined in (2.5) and for given $q \in \{1, \ldots, n\}$, set $r = q + k - 1$ if $q \leq n-k$ and $r = n$ otherwise. Let $E_r \in \mathbb{C}^{r\times r}$ be the matrix obtained from $E$ by Procedure 2.2.3 with $j = r-k$. Let, moreover, $X_q$ be the matrix from (2.2) partitioned as

$$X_q = \begin{bmatrix} Y_q \\ \lambda_q I_{n-r} \end{bmatrix}$$

where $Y_q = \begin{bmatrix} \Lambda_q \\ \lambda_q I_{r-q} \end{bmatrix}$

is an $r \times r$ matrix. Then,

$$\sigma(X_q + E) = \{\sigma(Y_q + E_r), \lambda_q\}.$$ 

Proof. The particular case of the smallest eigenvalue when $k = 1$ was treated in \[46\]. We look now into the general proof for all eigenvalues and for general $k$. 


Note that $Y_q$ is of size $r \times r$ containing the $r$ smallest eigenvalues of $X_q$ (counting multiplicities). By partitioning $W$ appropriately with $j = r - k$ according to Procedure 2.2.3 we get a unitary matrix $G$ such that

$$
G^* E G = \begin{bmatrix}
E_r \\
O_{n-r}
\end{bmatrix}.
$$

However, Procedure 2.2.3 also ensures that $G^* X_q G = X_q$ and, therefore,

$$
G^* (X_q + E) G = \begin{bmatrix}
Y_q + E_r \\
\lambda_q I_{n-r}
\end{bmatrix},
$$

proving the statement.

Remark 2.2.5. The matrix $R$ in (2.6) is of size $k \times k$. Consequently, if $q > n - k$ the complete matrix $X_q$ is required and, thus, $r = n$.

Remark 2.2.6. If any eigenvalue of $A$ has multiplicity larger than $k$ then by Proposition 2.2.1 the bounds are sharp.

We now provide an example detailing the computation of the lower bounds for a rank-2 perturbation of the matrix $A$ from Example 2.1.1

2.2.2 Example

Let $A \in \mathbb{C}^{5 \times 5}$ and $X_q \in \mathbb{C}^{5 \times 5}$ for $q \in \{1, \ldots, 5\}$ be the matrices from Example 2.1.1 and for $\Sigma = I_2$ and $W = [w|v] \in \mathbb{C}^{5 \times 2}$ let $E = W \Sigma W^*$ be a Hermitian matrix of rank-2. For this example we take the vectors $w = \begin{bmatrix}
1 & 1 & 1 & 1 & 1
\end{bmatrix}^*$ and $v = \begin{bmatrix}
1 & -1 & 1 & -1 & 1
\end{bmatrix}^*$. The $r \times r$ matrices $Y_q$ from (2.7) are displayed next

$$
Y_1 = \begin{bmatrix}
2 \\
2
\end{bmatrix}, \qquad Y_2 = \begin{bmatrix}
2 \\
3 \\
3
\end{bmatrix}, \qquad Y_3 = \begin{bmatrix}
2 \\
3 \\
3.1 \\
3.1
\end{bmatrix}.
$$
\[ Y_4 = \begin{bmatrix} 2 \\ 3 \\ 3.1 \\ 4 \\ 4 \end{bmatrix}, \quad \text{and} \quad Y_5 = A = \begin{bmatrix} 2 \\ 3 \\ 3.1 \\ 4 \\ 4.1 \end{bmatrix}. \]

We shall compute bounds for the eigenvalues of \( A + E \) using only \( Y_2 \) as the procedure for the remaining matrices is similar. In this case, \( r = 3 \) and following Procedure (2.2.3), we partition \( W \) as

\[
W = \begin{bmatrix} w_1 & v_1 \\ \tilde{w} & \tilde{v} \end{bmatrix}
\]

and compute a QR-decomposition of \( \begin{bmatrix} \tilde{w} & \tilde{v} \end{bmatrix} \)

\[
\begin{bmatrix} \tilde{w} & \tilde{v} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}.
\]

In this example, \( \tilde{w} \) and \( \tilde{v} \) have length 4 and are already orthogonal. Therefore, with the unitary matrix \( G = 1 \oplus Q \) we obtain

\[
G^* E G = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 4 \end{bmatrix} = \begin{bmatrix} E_3 \\ O_2 \end{bmatrix}.
\]

Finally, computing the eigenvalues of \( Y_2 + E_3 \) gives us lower bounds for the eigenvalues of \( A + E \). In Table 2.2 we display, for \( q \in \{1, \ldots, 5\} \) the eigenvalue of \( X_q + E \) from Example 2.2.2 computed using the approach of Theorem 2.2.2.
2.2. HERMITIAN RANK-$K$ PERTURBATIONS 35

\[ \lambda_1 \leq \cdots \leq \lambda_n \]

\[
\begin{array}{c|ccccc}
X_1 + E & 2 & 2 & 2 & 6 & 8 \\
X_2 + E & 2.2984 & 3 & 3 & 7 & 8.7016 \\
X_3 + E & 2.3247 & 3.0494 & 3.1 & 7.0506 & 8.7753 \\
X_4 + E & 2.3939 & 3.4384 & 3.5641 & 7.5616 & 9.1420 \\
A + E & 2.3981 & 3.4384 & 3.6149 & 7.5616 & 9.1870 \\
\end{array}
\]

Table 2.2: Eigenvalues of $X_q + E$ for $q \in \{1, \ldots, 5\}$ from Example 2.2.2 computed using the approach of Theorem 2.2.2.

2.2.3 Perturbations of non-diagonal matrices

For the problem we are considering, we could, without loss of generality, study diagonal matrices instead of arbitrary Hermitian matrices. In practice, however, since diagonalizations of large matrices are expensive it is wise to avoid them whenever possible. The bounds developed earlier can be computed without an explicit diagonalization. In fact, because for these perturbations, the knowledge of a few eigenvalues and eigenvectors suffices, the computational costs to determine the bounds are low.

Let $A$ be Hermitian with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$ and corresponding orthonormal eigenvectors $v_1, \ldots, v_n$. For given $j \in \{0, \ldots, n\}$, write $V_j$ for the matrix with columns $v_1, \ldots, v_j$ and $V_{n-j}$ for the matrix with columns $v_{j+1}, \ldots, v_n$ and let $V = (V_j V_{n-j})$. Now, for each matrix $X_q$ from (2.2), $V^* A V - X_q$ is positive semidefinite. Thus, $T_q = V X_q V^*$ is an approximation of $A$ from below, and for any Hermitian perturbation $E$ of $A$, the eigenvalues of $T_q + E$ are lower bounds for the corresponding eigenvalues of $A + E$. Of course, the eigenvalues of $T_q + E$ are the eigenvalues of $X_q + V^* E V$. Thus, with $E = W \Sigma W^*$ where $W$ is an $n \times k$ matrix with orthonormal columns, the problem that results is a rank-$k$ perturbation of a diagonal matrix.

Assume now that only the eigenpairs belonging to the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ with $1 \leq \ell < n$ are available. Then $V^* E V$ cannot be explicitly computed. But because $W$ has orthonormal columns, so has $V^* W$, and for each $j \in \{0, \ldots, n\}$ we find that

\[ I = (V^* W)^* (V^* W) = W^* V_j V_j^* W + W^* V_{n-j} V_{n-j}^* W. \]

Note that $Y^* Y = Z^* Z$ for matrices $Y$ and $Z$ of the same size, if and only if
Y = QZ for some unitary transformation Q. Therefore, let Y be any solution of the $k \times k$ matrix equation

$$Y^*Y = I - W^*V_jV_j^*W$$  \hspace{1cm} (2.9)$$

for instance, the $k \times k$ Cholesky factor of the right-hand side, padded with $n - j - k$ zero rows, then $Y = Q_{n-j}V_{n-j}^*W$ for some unitary $(n - j) \times (n - j)$ matrix $Q_{n-j}$. This yields

$$V^*EV = V^*W\Sigma W^*V = \begin{bmatrix} V_j^*W \\ V_{n-j}^*W \end{bmatrix} \Sigma \begin{bmatrix} W^*V_j & W^*V_{n-j} \end{bmatrix}$$

$$= \begin{bmatrix} I_j & Q_{n-j} \end{bmatrix}^* \begin{bmatrix} V_j^*W \\ Y \end{bmatrix} \Sigma \begin{bmatrix} W^*V_j & Y^* \end{bmatrix} \begin{bmatrix} I_j & Q_{n-j} \end{bmatrix}.$$  

If $j \leq \ell$, then the only unknown in the above expression is $Q_{n-j}$. But if $q \leq j$ then the eigenvalues of $X_q + V^*EV$ can still be computed, since $X_q$ commutes with the $n \times n$ unitary transformation, i.e., $X_q + V^*EV$ equals

$$\begin{bmatrix} I_j & Q_{n-j} \end{bmatrix}^* \left( X_q + \begin{bmatrix} V_j^*W \\ Y \end{bmatrix} \Sigma \begin{bmatrix} W^*V_j & Y^* \end{bmatrix} \right) \begin{bmatrix} I_j & Q_{n-j} \end{bmatrix}.$$  

What remains is a rank-$k$ perturbation of a diagonal matrix with known perturbation, which can be treated as in Theorem 2.2.4. In fact, if $Y$ is indeed upper triangular due to the choice of the Cholesky factor in and below (2.9), it is clear that the eigenvalues of $X_q + V^*EV$ that differ from the ones from $X_q$ are the eigenvalues of the leading principal submatrix of $X_q + V^*EV$ of size $(j + k) \times (j + k)$.

### 2.2.4 Alternative approximations

The matrices chosen in (2.2) are not the only approximations obtained from the eigenvalues of $A$ that approximate $A$ from below. In fact, we may construct other approximations satisfying the same property and, with that, still obtain lower bounds for the same eigenvalue of $A + E$.

In Example 2.1.1 improving the bounds was implicitly associated with a)
Knowing a larger number of eigenvalues of $A$; and b) Increasing the size of the small eigenvalue problem. Obtaining improved bounds without the knowledge of more eigenvalues and without increasing the size of the small eigenvalue problem may prove difficult. However, the construction of matrices leading to bounds of higher quality seems achievable without increasing $r$.

We turn to Example 2.1.1 and consider the following two matrices

\[
X_3 := \begin{bmatrix}
2 \\ 3 \\ 3.1I_3
\end{bmatrix}
\quad \text{and} \quad
X'_3 := \begin{bmatrix}
2 \\ 3I_2 \\ 4I_2
\end{bmatrix}
\]

(2.10)

together with Table 2.3 containing the respective eigenvalues after perturbation with $E$.

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$\lambda_3$</th>
<th>$\lambda_4$</th>
<th>$\lambda_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_3 + E$</td>
<td>2.1821</td>
<td>3.0250</td>
<td>3.1</td>
<td>3.1</td>
<td>7.8929</td>
</tr>
<tr>
<td>$X'_3 + E$</td>
<td>2.2143</td>
<td>3</td>
<td>3.4798</td>
<td>4</td>
<td>8.3059</td>
</tr>
<tr>
<td>$A + E$</td>
<td>2.2214</td>
<td>3.0475</td>
<td>3.5291</td>
<td>4.0519</td>
<td>8.3503</td>
</tr>
</tbody>
</table>

Table 2.3: Different lower bounds for the eigenvalues of $A + E$ given by the eigenvalues of $X_3 + E$ and $X'_3 + E$

For most eigenvalues, $X'_3$ offers tighter bounds than $X_3$ while still maintaining the size of the small eigenvalue problem unchanged. Note, moreover, that the error matrices satisfy $0.1 = \|S'_3\|_2 \leq \|S_3\|_2 = 1$ where $S'_3 = A - X'_3$ and $S_3 = A - X_3$. The idea here is to deviate from the use of multiple eigenvalues located at the bottom right positions of the matrix to the use of multiple eigenvalues located at any positions. Even if in the interior of the spectrum. The existence of eigenvalues with high multiplicity can be exploited for that effect.

In a way similar to the one used in Procedure 2.2.3, also here the matrix $E$ is transformed via a unitary matrix $G$ obtained from a direct sum of unitary submatrices. In turn, these are obtained from QR-decompositions of appropriate partitions of $W$. For instance, with the matrices from (2.10) and assuming that $E$ is of rank-one and that the focus is on the smallest eigenvalue of $A + E$, we may
partition $w$ as

$$w = \begin{bmatrix} w_1 \\ \vec{w}_2 \\ \vec{w}_3 \end{bmatrix}$$

and use the unitary matrix

$$G = \begin{bmatrix} 1 \\ Q_2 \\ Q_3 \end{bmatrix}$$

where $Q_2$ and $Q_3$ are such that $Q_2^* \vec{w}_2 = \|\vec{w}_2\|_2$ and $Q_3^* \vec{w}_3 = \|\vec{w}_3\|_2$ to obtain $G^*EG$. Note that $G^*EG$ is of the form

$$\begin{bmatrix} \star & \star & \star \\ \star & \star & \star \\ \star & \star & \star \end{bmatrix}$$

and thus additional permutations of rows and columns are needed to obtain $E_r$. It remains valid, however, that

$$\sigma(Y_q' + E_r) \subset \sigma(X_q' + E).$$

Depending on the eigenvalue for which an improved bound is required, the new matrices should satisfy two conditions: a) $\|S_q'\|_2 \leq \|S_q\|_2$; and b) if a bound for $\lambda_q(A + E)$ is to be determined, then $\lambda_q(X')$ should not belong to a cluster with other (different) eigenvalues. If these conditions are satisfied, both $X_q'$ and $X_q$ approximate $A$ from below. Moreover, under the assumption that all eigenvalues of $A$ are available, the computational effort required to determine the bounds is virtually the same.

**Remark 2.2.7.** Automatic clustering techniques can provide an optimal division of the spectrum. The computational effort required to execute such procedures is,
2.3 BORDERED DIAGONAL MATRICES

however, too high and would eliminate all advantage of not computing the exact eigenvalue. Our approach is intended to exploit available information.

2.3 Bordered diagonal matrices

We start this section by introducing the concept of Hermitian $k$-bordered diagonal matrices:

**Definition 2.3.1** (Hermitian $k$-bordered diagonal matrix). A Hermitian matrix is $k$-bordered diagonal if its entries are zero except for the main diagonal and $k$ rows and corresponding columns.

Formally, we pose no restriction on $k$ and, in particular, allow a diagonal matrix and a full matrix to be the 0 and $n - 1$-bordered diagonal matrices respectively. Moreover, owing to the invariance of the eigenvalues under similarity transformations, we opt to rearrange the matrix so that the nonzero rows and columns are the first $k$. We are thus interested in matrices of the form

$$
\begin{bmatrix}
C_1 & C_2 & \ldots & C_{n-k+1} \\
C_2^* & d_1 \\
\vdots & \ddots \\
C_{n-k+1}^* & \ldots & d_{n-k}
\end{bmatrix}
$$

(2.11)

where $C_1 \in \mathbb{C}^{k \times k}$ and for $j \in \{2, \ldots, n-1\}$, $C_j \in \mathbb{C}^{k}$ and where we assume that $d_1 \leq d_2 \leq \ldots \leq d_{n-k}$. For such matrices, the following result, showing a relation between Hermitian rank-$k$ perturbations of Hermitian matrices and $k$-bordered diagonal matrices, is known (see [83]) but we have not encountered it further in the literature dealing with bordered diagonal matrices.

**Lemma 2.3.2** ([83]). Let $n$ be an integer and let $A, E \in \mathbb{C}^{n \times n}$ be Hermitian matrices and assume further that $E$ is of rank-$k$. Then $A + E$ is unitarily similar to a Hermitian $k$-bordered diagonal matrix.
Proof. Let $U$ be unitary and such that

$$EU = U \begin{bmatrix} \Theta & \phantom{0} \\ 0 & O_{n-k} \end{bmatrix}$$

where $\Theta \in \mathbb{R}^{k \times k}$ is diagonal and contains the $k$ nonzero eigenvalues of $E$. We may apply the same unitary similarity transformation to $A$, yielding

$$AU = U \begin{bmatrix} M & B^* \\ B & T \end{bmatrix}$$

for certain $B \in \mathbb{C}^{(n-k) \times k}$ and for Hermitian $M \in \mathbb{C}^{k \times k}$ and $T \in \mathbb{C}^{(n-k) \times (n-k)}$. Now, let $V$ be unitary with $TV = VD$ where $D$ is the diagonal matrix of eigenvalues of $T$. Then, the unitary matrix

$$S = UG$$

where

$$G = \begin{bmatrix} I_k & \phantom{0} \\ V & \phantom{0} \end{bmatrix}$$

transforms $A + E$ into

$$S^*(A + E)S = \begin{bmatrix} M + \Theta & B^*V \\ V^*B & D \end{bmatrix}.$$ (2.12)

which is $k$-bordered diagonal.

Lemma 2.3.2 is of interest, not only on its own, but also because for $k = 1$, comparing the characteristic polynomial of $S^*(A + E)S$ with that of $A$, provides an alternative (and perhaps more widely known) proof to Proposition 2.2.1 (see [63] or [83]). The connections between rank-$k$ perturbations and bordered diagonal matrices, however, do not end here. The structure of bordered diagonal matrices allows for the computation of lower bounds for its eigenvalues using the principles discussed in Section 2.2.

Remark 2.3.3. From a perturbation point of view one could also treat a $k$-bordered diagonal matrix as a rank-$2k$ perturbation of a diagonal matrix. Such approach
would place us in the same setting of the previous section. However, the perturbation is indefinite and the element located at the tip of the arrow, \( c_1 \), is not necessarily smaller than \( d_1 \). A better approach is to treat the matrix on its own.

**Lemma 2.3.4.** Let

\[
A = \begin{bmatrix}
C_1 & C_2 & \ldots & C_{n-k+1} \\
C^*_2 & d_1 & & \\
& \ddots & \ddots & \\
C^*_{n-k+1} & & & d_{n-k}
\end{bmatrix}
\]

be a given \( k \)-bordered diagonal matrix in the conditions of (2.11). Moreover, for \( q \in \{1, \ldots, n-k\} \) let

\[
D_q := \begin{bmatrix}
d_1 \\
\ddots \\
d_q \\
d_q I_{n-q-k}
\end{bmatrix}
\text{ and set } A_q := \begin{bmatrix}
C_1 & C_2 & \ldots & C_{\ell} \\
C^*_2 & \ddots & \\
& \ddots & \ddots & \\
C^*_{\ell} & & & D_q
\end{bmatrix}
\]

where \( \ell = n-k+1 \). Then the eigenvalues of \( A_q \) are lower bounds for the eigenvalues of \( A \).

**Proof.** By direct computation it is easily seen that \( A - A_q \) is a real diagonal positive semi-definite matrix. \( \square \)

In view of Procedure 2.2.3 consider, now, Procedure 2.3.5 directed to the nonzero columns (and rows) of the matrix \( A_q \).

**Procedure 2.3.5.** Consider the 1-bordered diagonal matrix and let the elements of the nonzero column of \( A_q \) be the components of a vector \( c \), i.e., let \( c \) be the vector \( c = [c_2, c_3, \ldots, c_{n-1}, c_n]^* \). For \( j \in \{1, \ldots, n-2\} \) partition \( c \) as \( \tilde{c}_j \) containing the first \( j \) rows and the vector \( \tilde{c}_{n-1-j} \) with the remaining \( n-1-j \) rows:

\[
c = \begin{bmatrix}
\tilde{c}_j \\
\tilde{c}_{n-1-j}
\end{bmatrix}.
\]
Then, for a unitary matrix $Q$ such that

$$Q \begin{bmatrix} \|\tilde{c}_{n-1-j}\|_2 \\ 0 \end{bmatrix} = \tilde{c}_{n-1-j}$$

is a QR-decomposition of $\tilde{c}_{n-1-j}$ we have with $G = I_j \oplus Q$ that

$$G^* A_q G = \begin{bmatrix} c_1 & \cdots & c_j & \|\tilde{c}_{n-1-j}\|_2 & O \\ \vdots & & d_1 & & \vdots \\ c_j & & \ddots & & \vdots \\ \|\tilde{c}_{n-1-j}\|_2 & & & d_q & \vdots \\ O & & & d_q & d_q I_{n-r} \end{bmatrix} = \begin{bmatrix} Y_{j+1} \\ \vdots \\ d_q I_{n-r} \end{bmatrix}.$$

A procedure for $k$-bordered diagonal matrices with $k > 1$ can be described similarly.

As a consequence of Procedure 2.3.5, the eigenvalues of $A_q$ can be determined at the cost of an eigenvalue problem for a small $(j + 1) \times (j + 1)$ matrix.

**Theorem 2.3.6.** Let $1 \leq k \leq n$ and for $q \in \{1, \ldots, n - 2k\}$ set $r = q + 2k$. Then

$$\sigma(A_q) = \{\sigma(Y_r), d_q\}.$$

**Proof.** Applying Procedure 2.3.5 to an appropriate partitioning of the columns of $A_q$ and noting that the resulting matrix is block diagonal gives the statement. \qed

### 2.4 Singular Values

The Singular Value Decomposition (SVD) is a tool that has found several applications in areas such as Image Analysis [16] and Information Retrieval [8]. In both areas, the data is commonly represented in matrix form and the SVD is required. However, due to the dynamic nature of the data, new information can be added (or eventually removed) from the original set. In matrix terms this corresponds to adding or removing a column after which a new SVD is required [14, 16, 78]. In this section we are interested in two situations: a) lower bounds for the singular
values of rank-\(k\) perturbations of general \(n \times m\) matrices, \(B + K\); and b) lower bounds for the singular values of \(n \times m\) matrices when columns (or rows) are added.

### 2.4.1 Rank-\(k\) perturbations of general matrices

The study of singular values is typically done through the study of the eigenvalues of the Hermitian matrices resulting from right or left multiplication with the adjoint. We will follow the same approach here. The starting point is an \(n \times m\) matrix \(B\). Without loss of generality we assume that \(m \leq n\). This matrix will be perturbed with a rank-\(k\) matrix \(K\) and we are interested in the singular values of \(B + K\).

**Proposition 2.4.1.** Given the \(n \times m\) matrix \(B\), and a rank-\(k\) perturbation \(K\) of \(B\), the singular values of \(B + K\) are the square roots of the nonnegative eigenvalues of

\[
(B + K)^*(B + K) = B^*B + K^*V + V^*K,
\]

where \(V = B + \frac{1}{2}K\), which constitutes a Hermitian perturbation of rank at most \(2k\) of \(B^*B\).

**Proof.** Write \(V = B + \frac{1}{2}K\), then

\[
(B + K)^*(B + K) = B^*B + K^*B + B^*K + K^*K = B^*B + K^*(B + \frac{1}{2}K) + (B^* + \frac{1}{2}K^*)K.
\]

This proves the first part of the statement. To see that the perturbation has rank at most \(2k\), note that the rank of each of the products \(V^*K\) and \(K^*V\) is less than or equal to the rank of \(K\), which is \(k\).

The lower bounds for Hermitian rank-\(k\) perturbations of Hermitian matrices from Section 2.2 can now be applied to the specific perturbation of \(B^*B\) in (2.13). These bounds were in principle developed for perturbations \(E\) that were given in the form \(E = W\Sigma W^*\) with \(W^*W = I\). If the perturbation \(K\) of \(B\) is given as \(K = ZX^*\) for certain \(n \times k\) matrix \(Z\) and \(m \times k\) matrix \(X\), this can be exploited.
in the sense that, writing $S$ for the $m \times k$ matrix $V^*Z$,

$$K^*V + V^*K = X(Z^*V) + (V^*Z)X^*$$

$$= (S|X) \begin{bmatrix} I & \end{bmatrix} (S|X)^* = QR \begin{bmatrix} I & \end{bmatrix} R^*Q^*$$

where $(S|X) = QR$ is a QR-decomposition. Computing an eigenvalue decomposition of the $2k \times 2k$ Hermitian matrix

$$R \begin{bmatrix} I & \end{bmatrix} R^* = Y\Lambda Y^*$$

then leads to the desired format:

$$E = W\Sigma W^* \quad \text{with} \quad W = QY, \quad W^*W = I, \quad \text{and} \quad \Sigma = \Lambda. \quad (2.14)$$

Observe that the generic situation will be that $R$ is invertible. Then, by Sylvester’s Theorem for Inertia, $\Lambda$ has $k$ negative and $k$ positive eigenvalues, and thus, the perturbation of $B^*B$ is indefinite. This shows that one needs to take care when constructing lower bounds for the singular values of $B + K$. If $B^*B$ is replaced by an approximation from below based on some matrix $X_q$, adding the indefinite perturbation $K^*V + V^*K$ from (2.13) may lead to negative eigenvalues of $X_q + K^*V + V^*K$, rendering some of the lower bounds worthless in the context of singular values, being the square roots of eigenvalues.

**Remark 2.4.2.** If $K$ is given as its singular value decomposition $K = Z\Gamma X^*$ with $n \times k$ matrix $Z$ and $m \times k$ matrix $X$, both with orthonormal columns, and $\Gamma$ a real nonnegative $k \times k$ diagonal matrix, then by absorbing $\Gamma$ into $X^*$, $K$ will have the form used above.

**Proposition 2.4.3.** Given the $n \times m$ matrix $B$, and a rank-$k$ perturbation $K$ of $B$. Then the singular values of $B + K$ are the square roots of the $m$ largest nonnegative eigenvalues of the $n \times n$ matrix

$$(B + K)(B + K)^* = BB^* + KV^* + VK^*, \quad (2.15)$$
where \( V = B + \frac{1}{2}K \), which constitutes a Hermitian perturbation of rank at most \( 2k \) of \( BB^* \).

**Proof.** Write \( V = B + \frac{1}{2}K \), then

\[
(B + K)(B + K)^* = BB^* + BK^* + KB^* + KK^*
= BB^* + (B + \frac{1}{2}K)K^* + K(B^* + \frac{1}{2}K^*).
\]

This proves the first part of the statement. To see that the perturbation has rank at most \( 2k \), note that the rank of each of the products \( KV^* \) and \( VK^* \) is less than or equal to the rank of \( K \), which is \( k \).

In this approach, although seemingly similar to the previous, if \( m < n \), the \( n \times n \) matrix \( BB^* \) has \( (n - m) \) artificial zero eigenvalues. Although this is no restriction to apply the approach of Section 2.2, it may reduce the effectiveness of some of the bounds if.

The same distinction between the different ways to form the product of the matrix and its adjoint will be present in the following section.

**Addition of columns or rows**

Let \( k, n \) and \( m \) be integers with \( n > m \) and \( k \geq 1 \). Consider again the matrix \( B \in \mathbb{C}^{n \times m} \), and matrices \( C \in \mathbb{C}^{n \times k} \) and \( N \in \mathbb{C}^{m \times k} \). We aim to find lower bounds for the singular values of the matrices \( B_c \) and \( B_r \) defined by

\[
B_c = (B|C) \quad \text{and} \quad B_r = (B^*|N)^*;
\]

and for that we look at the eigenvalues of the \( (m + k) \times (m + k) \) Hermitian matrix \( B_c^*B_c \) and the \( n \times n \) Hermitian matrix \( B_cB_c^* \) and their counterparts for rows.
Case 1: Study of $B_c^* B_c$

The matrix $B_c^* B_c$ equals

$$
B_c^* B_c = \begin{bmatrix}
B^* B & B^* C \\
C^* B & C^* C
\end{bmatrix}.
$$

Now, let

$$
B = U \begin{bmatrix}
\Sigma_m \\
O
\end{bmatrix} W^*
$$

be an SVD of $B$, with $U \in \mathbb{C}^{n \times n}, W, \Sigma_m \in \mathbb{C}^{m \times m}$. Then with $G = W \oplus I_k$ we find that

$$
G^* (B_c^* B_c) G = \begin{bmatrix}
\Sigma_m^2 & Z \\
Z^* & C^* C
\end{bmatrix}
,$$

where $Z = W^* (B^* C)$.

This is (a block permutation of) a $k$-bordered diagonal matrix having no artificial zero eigenvalues. The results from Section 2.3 can now be applied to find lower bounds for its eigenvalues. Again, replacing the matrix by an approximation from below may result in a matrix with negative eigenvalues, thus one needs to be cautious.

Case 2: Study of $B_c B_c^*$

The matrix $B_c B_c^*$ equals

$$
B_c B_c^* = BB^* + CC^*.
$$

Transforming this matrix to the basis given by the columns of $U$ from (2.16) shows (with $C_1$ and $C_2$ chosen appropriately) that

$$
U^* (B_c B_c^*) U = \begin{bmatrix}
\Sigma_m^2 \\
O
\end{bmatrix} + \begin{bmatrix}
C_1 \\
C_2
\end{bmatrix} \begin{bmatrix}
C_1^* \\
C_2^*
\end{bmatrix}
,$$

with

$$
\begin{bmatrix}
C_1 \\
C_2
\end{bmatrix} = U^* C.
$$

Now, let

$$
C_2 = Q \begin{bmatrix}
R \\
O
\end{bmatrix}
$$
be a QR-decomposition of $C_2$, where $R$ is $k \times k$, then the at most $m + k$ nonzero eigenvalues of $U^*(B_cB_c^*)U$ are the eigenvalues of the $(m + k) \times (m + k)$ matrix

$$
\begin{bmatrix}
\Sigma_m^2 \\
O_k
\end{bmatrix} + 
\begin{bmatrix}
C_1 \\
R
\end{bmatrix} \begin{bmatrix}
C_1^* & R^*
\end{bmatrix}.
$$

This is a positive semi-definite perturbation of a diagonal matrix, and again the approach of Section 2.2 may be applied to find lower bounds for its eigenvalues. Since these lower bounds will be nonnegative, their square roots are well-defined, and thus valid lower bounds for the singular values of $B_c$ result.

Similar results can, of course, be obtained for the matrix $B_r$: the matrix $B_rB_r^*$ compares to $B_c^*B_c$ and $B_r^*B_r$ to $B_cB_c^*$.

## 2.5 Numerical illustrations

We now illustrate the techniques developed in the previous sections. Our focus is on the computation of lower bounds for the smallest eigenvalue of Hermitian rank-one perturbations of Hermitian matrices. For these matrices not only are the newly proposed bounds cheaply obtainable, but we are also able to compare them with the existing ones from Weyl and Ipsen and Nadler. A few illustrations are also given for rank-2 perturbations of Hermitian matrices and for the addition of columns to $n \times m$ matrices.

### Example 1

For our first numerical experiment we use the reaction-diffusion example given in §2.2 on $I = [0, 1]$:

$$
-u'' + su = \lambda u \quad \text{in } I, \quad \text{with } u(0) = u(1) = 0 \quad (2.17)
$$

where the function $s$ is a local reaction term defined in $(\frac{1}{2} - \xi, \frac{1}{2} + \xi)$ for $\xi \in [0, \frac{1}{2}]$ by

$$
s(x) = e^{-g(x)} > 0 \quad \text{where} \quad g(x) = \frac{\xi^2}{(x - (1/2 - \xi)) (x - (1/2 + \xi))} \quad (2.18)
$$
and zero elsewhere. Note that our choice of working in one dimension is not restrictive and similar situations can be set up in higher dimensions. Take $p \in \mathbb{N}$ and define $h = 1/p$ and $x_j = jh$ where $j \in \{0, \ldots, p\}$. The Finite Difference discretization of problem (2.17) on the grid points $x_0, \ldots, x_p$ yields two matrices. The first is the typical tridiagonal matrix with 2 in the diagonal and $-1$ in both co-diagonals which we denote by $A$. The eigenvalues and eigenvectors of $A$ are known. The matrix $E = \sum_j s(x_j)e_j e_j^*$, on the other hand, is a diagonal matrix whose rank is determined by $h$ and $\xi$ by truncating $2\xi/h$ at integer values.

We wish to determine lower bounds for the eigenvalues of $A + E$ or, on the basis of eigenvectors of $A$, those of $\Lambda + \hat{E}$ where the diagonal matrix $\Lambda$ and the unitary matrix $V$ are such that $A = V\Lambda V^*$ and

$$
\hat{E} = V^*EV = \sum_j s(x_j)v_j v_j^* \quad \text{with} \quad v_j = V^*e_j.
$$

We take $p = 32$, $\xi = 1/33$ such that $E$ from Equation (2.18) has rank-one. We compute lower bounds for the smallest eigenvalue of $A + E$ using the approaches of Section 2.2.1 and the technique in Section 2.2.3.

The top plot in Figure 2.1 shows the absolute value of the error in the lower bounds for the smallest eigenvalue of $A + E$. The lower bounds were obtained from approximations $X_q$ with $q$ ranging from 1 to $n$. Recall that each $X_q$ uses the $q$ smallest eigenvalues of $A$ with the largest of these having multiplicity at least $n - q + 1$. The errors for the first ten bounds are displayed again, for convenience, in the bottom plot of Figure 2.1. For notation simplicity we use in the label of the $y$-axis $\lambda_-$ to mean $\lambda_{\min}(A + E)$ and $\mu_-$ to denote $\lambda_{\min}(X_q + E)$.

Recall that, for $q = 1$, the bound obtained is the same as the one from Weyl's inequalities while for $q = 2$ the same from Ipsen and Nadler's results. Figure 2.1 shows that there is a slightly more pronounced error reduction at the beginning and at the end, while with the intermediate matrices we observe only minor accuracy improvements. This feature is shared by most of the matrices we have experimented with in this section. Furthermore, we observe from the bottom plot that although the second bound shows the best relative error reduction, it is still advantageous to use larger matrices if more accuracy is required. We expect this
2.5. NUMERICAL ILLUSTRATIONS

Figure 2.1: Error in the lower bound (and detail for $q \leq 10$) of the smallest eigenvalue of the perturbed matrix $A$ resulting from the Finite Difference discretization ($n = 32$, $\xi = 1/33$) of Problem (2.17) with $E$ a rank-one diagonal matrix resulting from Equation (2.18).

to be more pronounced with larger matrices.

Example 2 For our second example we chose a larger Hermitian matrix $A$ of dimension $n = 1600$ resulting from the standard finite difference discretization of the Poisson equation in a square domain using a $40 \times 40$ grid. We perturb this matrix with a random Hermitian matrix of rank-one. The absolute value of the errors for the lower bounds obtained with different approximating matrices $X_q$ is represented in the two plots of Figure 2.2. For the same problem, Table 2.4 shows the approximate values of the error for the first six lower bounds in comparison to the ones for the matrix of Example 3 to be discussed further on.

We can observe from Table 2.4 and from the second plot in Figure 2.2 that the initial error reduction is, in this case, more pronounced. Consequently, and at a cost of a $6 \times 6$ eigenvalue problem ($q = 5$), Weyl’s bounds and those from [46] can be improved. The additional cost is modest while the absolute value of the
error decreases from $O(10^{-2})$ to $O(10^{-6})$. Another interesting observation is the presence of line segments parallel to the $x$-axis between the values $2 – 3$, $5 – 6$, $7 – 8$ and $9 – 10$. They correspond to multiple eigenvalues (in this case of multiplicity 2) of the matrix $A$. Note that, according to Proposition 2.2.1 and Section 2.2 these lines should still be visible when perturbing the same matrix with a rank-2 perturbation because the two consecutive matrices $X_q$ are equal. We illustrate this in Figure 2.3 where we represent the absolute value of the error for the lower bounds of a random Hermitian rank-2 perturbation of the same matrix $A$.

\[
\begin{array}{ccc}
q & |\lambda_- - \mu_-| & q & |\lambda_- - \mu_-| \\
1 & 1.756e^{-2} & 1 & 0.2959e^{-3} \\
2 & 8.544e^{-3} & 2 & 0.2917e^{-3} \\
3 & 8.544e^{-3} & 3 & 0.2917e^{-3} \\
4 & 7.572e^{-5} & 4 & 0.2838e^{-3} \\
5 & 5.586e^{-6} & 5 & 0.2702e^{-3} \\
6 & 5.586e^{-6} & 6 & 0.2638e^{-3} \\
\ldots & \ldots & \ldots & \ldots \\
\end{array}
\]

Table 2.4: Error in $\mu_- := \lambda_{min}(Y_r + \hat{E})$ for $1 \leq r \leq n$ for a first few lower bounds for the matrix obtained from the discretization of Poisson equation on a $40 \times 40$ grid (left) and Sherman1 from Matrix Market (right). The perturbations are random rank-one matrices.

**Example 3** For the third example we retrieve the matrix Sherman1 from the Matrix Market repository [1]. This is a real symmetric matrix of size $n = 1000$ with all eigenvalues situated on the negative real line. The perturbation was again a random Hermitian matrix of rank-one. We observe both from the second plot of Figure 2.4 as well as from the second column of Table 2.4 that the results for this matrix are not as good as the ones from Example 2. However, this example is interesting for just that. The stagnation period between $q = 200$ and $q = 500$ is due to the existence of an eigenvalue of multiplicity 300. We have mentioned in Section 2.2.4 that multiple eigenvalues are the most simple clusters and that these can be exploited to obtain improved bounds without increasing the size of the approximating matrix.

We illustrate this situation in Figure 2.4 where to construct the approximations of the matrix $A$ we make use of the knowledge of an eigenvalue with large
2.5. NUMERICAL ILLUSTRATIONS

Figure 2.2: Error in the lower bound (and detail for $q \leq 10$) of the smallest eigenvalue of a random rank-one perturbation $E$ of the matrix $A$ obtained from the discretization of Poisson equation on a 40 $\times$ 40 grid.

Figure 2.5 plots the absolute value of the error in the lower bounds for the smallest singular value of $B_c$ given by different approximating matrices. In the bottom plot we display the first twenty elements. Once again, a few small eigenproblems suffice to obtain reasonable approximations for the singular values. For this matrix and due to the fast error reduction stage at the beginning, using more than the second
lower bounds is certainly worthwhile.

2.6 Final notes

The previous sections have dealt with Hermitian matrices. Some of the theory developed, however, can be used with matrices other than Hermitian. In the next few pages we comment on examples of these situations.

Remark 2.6.1. If $A$ and $E$ are Hermitian matrices, $iA$ and $iE$ are skew-Hermitian and we can use the previous theory to study skew-Hermitian perturbations of skew-Hermitian matrices. In fact, for $\theta \in [0, 2\pi)$ this is true for any matrices $e^{2\pi i \theta} A$ and $e^{2\pi i \theta} E$. In the next chapter we investigate a more general class of normal matrices of which these are a particular case.

2.6.1 Non-Hermitian perturbations of Hermitian matrices

The notions of lower or upper bounds for eigenvalues lose their meaning for most non-Hermitian matrices. Some of the techniques described earlier, however, are

Figure 2.3: Error in the lower bound (and detail for $q \leq 10$) of the smallest eigenvalue of the perturbed matrix Poisson with $E$ a rank-2 random matrix.
Figure 2.4: Error in the lower bound of the smallest eigenvalue of the perturbed matrix $A$ for the matrix Sherman1 when a multiple eigenvalue is clustered (below) and comparison with the naive approach (above).

Figure 2.5: Error in the lower bound (and detail for $q \leq 20$) of the smallest singular value of $B_c$ where $B$ is the matrix ILLC1850 of dimensions $1850 \times 712$ from the Matrix Market and $c$ is a random $n$-dimensional vector.
still valid. The next result shows that when a Hermitian matrix $A$ is perturbed by a non-Hermitian matrix $E$, we can still use similarity transformations to reduce the size of the eigenvalue problem for $A + E$. We begin by presenting a useful procedure.

**Procedure 2.6.2.** Given $v, w \in \mathbb{C}^n$ define the matrix $Z = \begin{bmatrix} v & w \end{bmatrix}$. Partition each column of $Z$ in $p$ blocks (which, for notational simplicity, we assume now to be two) as

$$Z = \begin{bmatrix} \vec{v}_1 & \vec{w}_1 \\ \vec{v}_2 & \vec{w}_2 \end{bmatrix}.$$

For $p \in \{1, 2\}$ let $Q_p$ and $R_p$ be such that

$$Q_p \begin{bmatrix} R_p \\ O \end{bmatrix}$$

is a QR-decomposition of each partition $[\vec{v}_p \ \vec{w}_p]$ and denote by $Q$ and $R$ the matrices

$$Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} R_1^* & O & R_2^* & O \end{bmatrix}^*$$

respectively. Note now that

$$Q^*vw^*Q = \begin{bmatrix} \ast \\ 0_{\ell_1} & \ast \\ \ast & 0_{\ell_1-1} & \ast & \ast \\ 0_{\ell_2} & \ast & 0_{\ell_2-1} \end{bmatrix} = zy^*.$$

We now use Procedure 2.6.2 to prove the following theorem.

**Theorem 2.6.3.** Given $w, v \in \mathbb{C}^n$ set $E := vw^*$. Let $X \in \mathbb{R}^{n \times n}$ be diagonal with exactly $t$ distinct eigenvalues denoted by $\lambda_p$ ($1 \leq p \leq t$) with multiplicity $m_p$. Let $z$ and $y$ be obtained from $v$ and $w$ using Procedure 2.6.2 with $p = t$ and $v_p, w_p \in \mathbb{C}^{m_p}$. Then the eigenvalues of $X + E$ are the eigenvalues of $X$, of which double copies
of multiple eigenvalues are replaced by the eigenvalues of the $2t \times 2t$ matrix

$$Y = \begin{bmatrix} \lambda_1 I_2 & \cdots & \lambda_1 I_2 \\ \vdots & \ddots & \vdots \\ \lambda_2 I_2 & \cdots & \lambda_2 I_2 \end{bmatrix} + zy^*.$$  

Moreover,

$$\{\lambda_1, \ldots, \lambda_t\} \subset \sigma(Y).$$

Proof. We give the details of the proof for $t = 2$ after which the proof for $t > 2$ should be clear. Let

$$X = \begin{bmatrix} \lambda_1 I_{m_1} \\ \lambda_2 I_{m_2} \end{bmatrix}$$

where for $m_p > 1$ is the multiplicity of $\lambda_1$ and $m_2 > 1$ is the multiplicity of $\lambda_2$ and $m_1 + m_2 = n$. Using Procedure 2.6.2 we obtain a unitary matrix $Q$ such that $Q^*vw^*Q = zy^*$ and $Q^*XQ = X$. Therefore,

$$Q^*(X + E)Q = X + Q^*vw^*Q = X + zy^*.$$  

By appropriate permutations we may write $X + zy^*$ as $Y \oplus \lambda_1 I_{m_1-2} \oplus \lambda_2 I_{m_2-2}$ where

$$Y = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} + \begin{bmatrix} * & * & * & * \\ * & 0 & * & * \\ 0 & * & * & * \end{bmatrix}.$$  

Note that $Y$ can be seen as the sum of an upper triangular matrix with a Hermitian rank-one and thus the eigenvalues on the two last diagonal positions of the diagonal matrix are still eigenvalues of $Y$. Therefore, $m_p - 1$ copies of the eigenvalues of $X$ are still eigenvalues of $X + E$. This proves the statement for $t = 2$. For $t > 2$ the proof is similar. \hfill \Box

When compared with Proposition 2.2.1 the consequences of the loss of symmetry are clear. The resulting matrix is now of size $2t \times 2t$ instead of size $t \times t$ for the Hermitian case.
2.6.2 Non-Hermitian \( k \)-bordered diagonal matrices

In Lemma 2.3.2 we have related Hermitian rank-\( k \) perturbations of Hermitian matrices to Hermitian \( k \)-bordered diagonal matrices. When the perturbation is non-Hermitian, however, we are still able to show a similar result.

**Lemma 2.6.4.** Let \( n \) be an integer and let \( A \in \mathbb{C}^{n \times n} \) be Hermitian and \( E \in \mathbb{C}^{n \times n} \) be a non-Hermitian matrix. If \( E \) is of rank-\( k \) then \( A + E \) is unitarily similar to a non-Hermitian \( k \)-bordered diagonal matrix.

**Proof.** Let \( U \in \mathbb{C}^{n \times n} \), \( \Theta \in \mathbb{C}^{k \times k} \) and \( L \in \mathbb{C}^{(n-k) \times k} \) be such that

\[
U^*EU = \begin{bmatrix} \Theta & L^* \\ O_{n-k} & \end{bmatrix}
\]

is a Schur decomposition of \( E \) and consider the matrix

\[
U^*AU = \begin{bmatrix} M & B^* \\ B & T \end{bmatrix},
\]

for certain \( M \in \mathbb{C}^{k \times k} \), \( B \in \mathbb{C}^{(n-k) \times k} \) and (Hermitian) \( T \in \mathbb{C}^{(n-k) \times (n-k)} \). Now, let \( V \) be unitary and satisfy \( V^*TV = D \) with \( D \) the diagonal matrix of eigenvalues of \( T \). Defining the unitary matrices

\[
G = \begin{bmatrix} I_k \\ V \end{bmatrix} \quad \text{and} \quad S = UG,
\]

\( A + E \) can now be transformed as

\[
S^*(A + E)S = \begin{bmatrix} M + \Theta & (B^* + L^*)V \\ V^*B & D \end{bmatrix}.
\] (2.19)

which is \( k \)-bordered diagonal.

In the same way, it is possible (in the presence of a diagonal element of large multiplicity) to reduce the size of a non-Hermitian \( k \)-bordered diagonal eigenproblem by unitary similarity transformations. This is described in Procedure 2.6.5.
2.6. FINAL NOTES

Procedure 2.6.5. Consider the non-Hermitian 1-bordered diagonal matrix $A \in \mathbb{C}^{n \times n}$

$$A = \begin{bmatrix} a_1 & w_2 & \ldots & w_q & \tilde{w}^* \\ v_1 & d_1 \\ \vdots & \ddots \\ v_q & d_q \\ \tilde{v} & d_{q+1}I_{n-q-1} \end{bmatrix}$$

where $\tilde{v} = [v_{q+1}, \ldots, v_{n-1}]^*$ and $\tilde{w} = [w_{q+1}, \ldots, w_{n-1}]^*$. Let $QR$ be a QR-decomposition of $Z = [\tilde{v} \quad \tilde{w}]$. The matrix $G = I_q \oplus Q$ is then unitary and transforms $A$ into

$$G^*AG = \begin{bmatrix} D_1 \\ d_{q+1}I_{n-q-3} \end{bmatrix}$$

where

$$D_1 = \begin{bmatrix} a_1 & w_2 & \ldots & w_q & * & * \\ v_1 & d_1 \\ \vdots & \ddots \\ v_q & d_q \\ * & d_{q+1} \\ 0 & d_{q+1} \end{bmatrix}.$$

Therefore, the eigenvalues of $A$ are the eigenvalues of $D_1$ together with $d_{q+1}$ with multiplicity $n - q - 3$

Even though the eigenvalues of $X_q + E$ no longer bound the eigenvalues of $A + E$, they can still be considered approximations for them. We end this chapter with an illustration of this situation.

2.6.3 Example

Let $w = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}^*$ and $v = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 \end{bmatrix}^*$ and consider the matrix $E = ww^*$ as well as the matrix $A$ from Example 2.1.1. The approximate eigenvalues of $X_3 + E$ and $A + E$ are displayed in Table 2.5.

Note that by applying the Bauer-Fike Theorem (Theorem 1.1.1) to the matrix $(A + E) + (X_3 - A) = X_3 + E$ we would only expect that for at least one eigenvalue
Table 2.5: Eigenvalues of $X_3 + E$ and $A + E$ where $E$ is a non-Hermitian perturbation.

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$\lambda_3$</th>
<th>$\lambda_4$</th>
<th>$\lambda_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_3 + E$</td>
<td>2.76-0.34i</td>
<td>2.76+0.34i</td>
<td>3.1</td>
<td>3.1</td>
<td>3.58</td>
</tr>
<tr>
<td>$A + E$</td>
<td>2.77-0.36i</td>
<td>2.77+0.36i</td>
<td>3.59-0.25i</td>
<td>3.59+0.25i</td>
<td>4.48</td>
</tr>
</tbody>
</table>

\[
|\lambda - \mu| \leq \kappa_2(V)\|X_3 - A\|_2 \approx 12
\]

where $V$ diagonalizes $A + E$ and $\mu$ is an eigenvalue of $X_3 + E$. This shows the potential of approximating the eigenvalues of $A$ in this manner.