On the condition numbers of a multiple eigenvalue of a generalized eigenvalue problem

Nakatsukasa, Yuji

2011

MIMS EPrint: **2011.85**

Manchester Institute for Mathematical Sciences

School of Mathematics

The University of Manchester

On the condition numbers of a multiple eigenvalue of a generalized eigenvalue problem

Yuji Nakatsukasa

Abstract For standard eigenvalue problems, closed-form expressions for the condition numbers of a multiple eigenvalue are known. In particular, they are uniformly 1 in the Hermitian case and generally take different values in the non-Hermitian case. We consider the generalized eigenvalue problem and identify the condition numbers. Our main result is that a multiple eigenvalue generally has multiple condition numbers, even in the Hermitian definite case. The condition numbers are characterized in terms of the singular values of the outer product of the corresponding left and right eigenvectors.

1 Introduction

This paper is concerned with the sensitivities of a multiple eigenvalue of generalized eigenvalue problems $Ax = \lambda Bx$ under normwise additive perturbations. When perturbations are introduced in the matrices, a multiple eigenvalue λ_0 of multiplicity r generally splits into r simple eigenvalues. We are interested in the sensitivity of each of the r eigenvalues in the absolute (as opposed to relative) sense.

The condition number $\kappa(A, \lambda_0)$ of a simple eigenvalue λ_0 of an $n \times n$ matrix A is defined as follows: denoting by $\lambda_0 + \Delta(\epsilon, E)$ the eigenvalue of $A + \epsilon E$ with $\|E\|_2 = 1$ such that $\lim_{\epsilon \to 0} \Delta(\epsilon, E) = 0$,

$$\kappa(A, \lambda_0) = \lim_{\epsilon \to 0} \sup_{\|E\|_2 = 1} \frac{|\Delta(\epsilon, E)|}{|\epsilon|}.$$

Y. Nakatsukasa
School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK. E-mail: yuji.nakatsukasa@manchester.ac.uk

In words, $\kappa(A, \lambda_0)$ measures by how much small changes in A can be magnified in the eigenvalue λ_0, in the first order sense. It is well-known (e.g. [4, p. 323]) that $\kappa(A, \lambda_0) = 1/|y^H x|$, where y and x are the left and right eigenvectors corresponding to λ, normalized to have unit 2-norms.

Sun [16] extends the notion of condition numbers to a multiple eigenvalue. Supposing that λ_0 is a nondefective multiple eigenvalue of A of multiplicity r, he defines r condition numbers $\kappa_i(A, \lambda_0)$ for $i = 1, \ldots, r$ by

$$\kappa_i(A, \lambda_0) = \lim_{\epsilon \to 0} \sup_{\|E\|_2=1} \frac{|\Delta_i(\epsilon, E)|}{|\epsilon|},$$

where $\lambda_0 + \Delta_i(\epsilon, E)$ for $i = 1, \ldots, r$ are (generally distinct) eigenvalues of $A + \epsilon E$, such that $|\Delta_1(\epsilon, E)| \geq |\Delta_2(\epsilon, E)| \geq \cdots \geq |\Delta_r(\epsilon, E)|$ and $\lim_{\epsilon \to 0} \Delta_i(\epsilon, E) = 0$ for all i. $\Delta_i(\epsilon, E)$ can be a discontinous function of ϵ, but $|\Delta_i(\epsilon, E)|$ cannot. Sun derives the closed-form expression

$$\kappa_i(A, \lambda_0) = \left(\prod_{j=1}^{i} c_j(A, \lambda_0) \right)^{1/i} \quad \text{for} \quad i = 1, \ldots, r, \tag{1}$$

where $c_i(A, \lambda_0)$ are the secants of the canonical angles between the left and right invariant subspaces corresponding to the multiple eigenvalue λ_0. When A is non-Hermitian $c_i(A, \lambda_0)$ generally take different values for different i, hence so do $\kappa_i(A, \lambda_0)$ and (1) shows that a multiple eigenvalue has multiple condition numbers.

In contrast, in the Hermitian case we have $c_i(A, \lambda_0) \equiv 1$, so $\kappa_i(A, \lambda_0) = 1$ for $i = 1, \ldots, r$, regardless of whether the perturbation is Hermitian or non-Hermitian. Hence (1) also shows the well-known fact that a multiple eigenvalue of a Hermitian matrix always has the same condition number 1 under general normwise perturbations.

The purpose of this paper is to extend the above results to generalized eigenvalue problems $Ax = \lambda Bx$. We follow the previous studies and define the condition numbers $\kappa_i(A, B, \tau, \lambda_0)$ of a multiple eigenvalue λ_0 by

$$\kappa_i(A, B, \tau, \lambda_0) = \lim_{\epsilon \to 0} \sup_{\|E\|_2=1, \|F\|_2=\tau} \frac{|\Delta_i(\epsilon, E, F)|}{|\epsilon|}, \tag{2}$$

where $\lambda_0 + \Delta_i(\epsilon, E, F)$ for $i = 1, \ldots, r$ are the r eigenvalues of $(A + \epsilon E, B + \epsilon F)$, such that $|\Delta_1(\epsilon, E, F)| \geq \cdots \geq |\Delta_r(\epsilon, E, F)|$ and $\lim_{\epsilon \to 0} \Delta_i(\epsilon, E, F) = 0$ for all i. Here τ is a positive constant that allows for the case where perturbations in A and B occur in different magnitudes, which is a notion adopted for example in [5].

Since a standard non-Hermitian eigenvalue problem can be regarded as a special case $(B = I, \tau = 0)$ of the generalized non-Hermitian eigenvalue problem $Ax = \lambda Bx$, it is clear that a multiple eigenvalue in the latter case must have multiple condition numbers. On the other hand, in the important case of the generalized Hermitian definite pair (where A and B are Hermitian and B is positive definite), it is not obvious whether or not the condition

numbers $\kappa_i(A, B, \tau, \lambda_0)$ for $i = 1, \ldots, r$ take different values. In this paper we derive closed-form expressions for $\kappa_i(A, B, \tau, \lambda_0)$ that reveal they generally do take r different values. We shall see that there are two sources for this difference in conditioning, namely the difference between the left and right eigenvectors (as present in non-Hermitian standard eigenproblems), and the fact that the B-orthonormal eigenvectors have 2-norms that are different from each other when $B \neq I$.

It is important to note that in the Hermitian definite case an alternative, natural choice of metric can be based on the B-based inner product $(x, y)_B = x^H By$ instead of the standard inner product $(x, y) = x^H y$. An analysis using the B-based inner product leads to the standard eigenvalue problem for the Hermitian matrix $B^{-1/2}AB^{-1/2}$, so in this inner product all the condition numbers are the same. Our discussion in this paper assumes the use of the standard inner product. We also note that the chordal metric $\mathrm{chord}(a, b) = |a-b|/\sqrt{(1 + |a|^2)(1 + |b|^2)}$ is commonly used when dealing with generalized eigenvalue problems [4, p. 378],[13, p. 283]. The condition numbers $\kappa_i(A, B, \tau, \lambda_0)$ in this paper can be recast in terms of the chordal metric by the scaling $\kappa_{i,\mathrm{chord}}(A, B, \tau, \lambda_0) \leftarrow \kappa_i(A, B, \tau, \lambda_0)/(1 + |\lambda_0|^2)$.

The condition numbers κ_i for $i = 1, \ldots, r$ of a nondefective finite multiple eigenvalue λ_0 in four situations are summarized in Table 1 (in which both the matrices A (and B) and the perturbations E (and F) are Hermitian in the middle column and non-Hermitian in the right column), expressed in terms of $\sigma_1, \ldots, \sigma_r$, the r positive singular values of $X_1 Y_1^H$, where X_1 and Y_1 are $n \times r$ matrices whose columns represent the right and left eigenvectors corresponding to λ_0 respectively, such that $(Y_1^H A X_1, Y_1^H B X_1) = (\lambda_0 I_r, I_r)$. In the standard case $B = I$ we have $\sigma_i = c_i(A, \lambda_0)$ for $i = 1, \ldots, r$. The contribution of this paper is that we fill in the second row of Table 1, that is, we identify the condition numbers of a multiple eigenvalue in a generalized eigenvalue problem, both for the Hermitian definite case and the non-Hermitian case. Note that for the Hermitian definite case, the perturbed pair is also definite if ϵF is sufficiently small ($\|\epsilon F\|_2 \leq \sigma_{\min}(B)$ is a sufficient condition, where $\sigma_{\min}(B)$ is B's smallest singular value), which is true in the limit $\epsilon \to 0$.

Table 1 Summary of condition numbers κ_i of a finite, nondefective multiple eigenvalue λ_0 for $i = 1, \ldots, r$.

	Hermitian	Non-Hermitian
$Ax = \lambda x$	1	$\left(\prod_{j=1}^{i} \sigma_j \right)^{1/i}$
$Ax = \lambda Bx$	$(1 + \tau|\lambda_0|) \min_{1 \leq j \leq i} \sqrt{\sigma_j \sigma_{i-j+1}}$	$(1 + \tau|\lambda_0|) \left(\prod_{j=1}^{i} \sigma_j \right)^{1/i}$

There are a number of related studies in the literature. The authors in [6] investigate the Hölder condition number, which is essentially κ_1 in our terminology when λ_0 is nondefective. The focus of [6] is the effect of the structure

of the perturbation on the Hölder condition number, and in section 4.1 we discuss how our results are related to those in [6].

An observation that a multiple eigenvalue of a generalized eigenvalue problem has different sensitivities under perturbations was first made in [13, p. 300], which mentions that a multiple eigenvalue of a pair such as $A = \left[\begin{smallmatrix} 2000 & 0 \\ 0 & 2 \end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix} 1000 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ tends to behave differently under perturbations in A and B. We note that as shown in [1,3,17] for Hermitian definite pairs, small Hermitian componentwise relative changes in A and B can introduce only small relative perturbation to any eigenvalue, and it is easy to see the two eigenvalues of the above pair (A, B) have similar behaviors under relative perturbations. However, in terms of "standard" normwise perturbation, that is, when (A, B) is perturbed to $(A + \epsilon E, B + \epsilon F)$ under $\|E\|_2 \leq 1, \|F\|_2 \leq \tau$ and $\epsilon > 0$, a multiple eigenvalue can exhibit different behaviors, for both Hermitian and non-Hermitian perturbations. The papers [7,9] consider the Hermitian definite case and give explanations for this behavior, presenting r different perturbation bounds for λ_0 under Hermitian perturbations of finite norm. The approach of this paper is different in that our focus is on the condition numbers, which are attainable perturbation bounds in the first order sense in the limit $\epsilon \to 0$. The bounds in [7,9] are valid for non-asymptotic $\epsilon > 0$ but are less tight (generally not attainable) when $\epsilon \to 0$. In addition, we consider both Hermitian and non-Hermitian perturbations.

Our arguments closely follow that of [16], in which the condition numbers are called *worst-case* condition numbers, to emphasize the difference from the *typical-case* condition numbers, as presented in [14]. In this sense, our results should also be regarded as worst-case condition numbers, in that κ_i are the largest attainable bounds in the first order sense. Experiments suggest that these bounds are not likely to be attained in practice for randomly generated perturbations, especially for large i (see the example in section 4.1).

The structure of this paper is as follows. First in section 2 we establish characterizations of the condition numbers $\kappa_i(A, B, \tau, \lambda_0)$ as defined in (2). In section 3 we treat the Hermitian definite case and show that a multiple eigenvalue indeed takes multiple condition numbers in generalized eigenvalue problems. Section 4 treats the general non-Hermitian case, in which we give simple numerical examples to illustrate our results. We also discuss the case of an infinite and defective multiple eigenvalue.

Notation: $\sigma_i(X)$ denotes the ith largest singular value of a rectangular matrix X. $\lambda_i(A)$ denotes the ith largest eigenvalue of A in magnitude. We use only the matrix spectral norm $\| \cdot \|_2$, so that $\|A\|_2 = \sigma_1(A)$. We use MATLAB notation, in which $V(i, j : k)$ with $j \leq k$ denotes a row vector consisting of the jth to kth elements of the ith row of V. I_k is the identity matrix of order k.

2 Characterizations of the condition numbers

To analyze the condition numbers in (2) we need the first order expansion for λ_0. The first order perturbation expansion for a simple eigenvalue is a

well-known result [12,13], and that for a multiple eigenvalue is also studied in [15] for standard eigenvalue problems and in [2] for generalized eigenvalue problems, including singular matrix pairs. For an $n \times n$ matrix pair (A, B), suppose that λ_0 is a nondefective finite multiple eigenvalue (we discuss the infinite and defective cases later in section 4.3) of multiplicity r, so that there exist nonsingular matrices $X = [X_1, X_2]$ and $Y = [Y_1, Y_2]$ with $X_1, Y_1 \in \mathbb{C}^{n \times r}$ such that

$$Y^H A X = \begin{bmatrix} \lambda_0 I_r & 0 \\ 0 & J_A \end{bmatrix}, \quad Y^H B X = \begin{bmatrix} I_r & 0 \\ 0 & J_B \end{bmatrix}. \tag{3}$$

Here the spectrum of the pair (J_A, J_B) does not contain λ_0. Then, the pencil $(A + \epsilon E, B + \epsilon F)$ has eigenvalues $\widehat{\lambda}_1, \widehat{\lambda}_2, \ldots, \widehat{\lambda}_r$ admitting the first order expansion [6,8]

$$\widehat{\lambda}_i = \lambda_0 + \lambda_i(Y_1^H(E - \lambda_0 F)X_1)\epsilon + o(\epsilon), \quad i = 1, 2, \ldots, r, \tag{4}$$

where $\lambda_i(Y_1^H(E - \lambda_0 F)X_1)$ are the eigenvalues of $Y_1^H(E - \lambda_0 F)X_1$ for $i = 1, \ldots, r$. In light of (4) and the definition (2), we obtain the following characterization of $\kappa_i(A, B, \tau, \lambda_0)$.

Lemma 1 *Let (A, B) be an $n \times n$ matrix pair that has the decomposition* (3), *so that λ_0 is a multiple eigenvalue of multiplicity r. Then, $\kappa_i(A, B, \tau, \lambda_0)$ in* (2) *can be expressed as*

$$\kappa_i(A, B, \tau, \lambda_0) \equiv \sup_{\|E\|_2 \leq 1, \|F\|_2 \leq \tau} |\lambda_i(Y_1^H(E - \lambda_0 F)X_1)|, \quad i = 1, \ldots, r, \tag{5}$$

where the eigenvalues $\lambda_i(Y_1^H(E - \lambda_0 F)X_1)$ are ordered by decreasing absolute value.

We note that the matrices E, F that attain the supremum in (5) are different for each value of i, that is, the supremums generally cannot be attained simultaneously. This fact is also noted in [15] for standard eigenvalue problems.

We now derive another characterization of $\kappa_i(A, B, \tau, \lambda_0)$, which can be regarded as a generalization of the fact $\kappa_i(A, \lambda_0) = \sup_{\|\widetilde{E}\|_2 \leq 1} |\lambda_i(\Sigma \widetilde{E})|$ for $i = 1, \ldots, r$ for the standard case $B = I$ [16, Thm. 2.1], where $\widetilde{E}, \Sigma \in \mathbb{C}^{r \times r}$ and $\Sigma = \text{diag}(c_i(A, \lambda_0))$. We note that the secants of the canonical angles have the expression $c_i(A, \lambda_0) = \sigma_i(X_1 Y_1^H)$ for $i = 1, \ldots, r$ [10,14]. In this paper we use the quantity $\sigma_i(X_1 Y_1^H)$ instead of $c_i(A, \lambda_0)$ to express the condition numbers, because it allows us to treat generalized eigenvalue problems in a uniform way.

We use the proof of Theorem 2.1 in [16], whose crucial identity is

$$|\lambda_i(Y_1^H(E - \lambda_0 F)X_1)| = |\lambda_i((E - \lambda_0 F)X_1 Y_1^H)| = |\lambda_i((E - \lambda_0 F)U\Sigma V^H)|$$
$$= |\lambda_i(V^H(E - \lambda_0 F)U\Sigma)| \quad \text{for} \quad i = 1, \ldots, r,$$

where $X_1 Y_1^H = U\Sigma V^H$ is the "thin" SVD, in which U and V have $r(\leq n)$ columns [4, p. 72]. Here, to get the first and last equalities we used the fact [13, p. 27] that for general $X \in \mathbb{C}^{n \times m}$ and $Y \in \mathbb{C}^{m \times n}$, the nonzero eigenvalues of

XY and those of YX are the same. Since V and U have orthonormal columns and E, F can be arbitrary matrices with $\|E\|_2 \leq 1, \|F\|_2 \leq \tau$, it follows that $\widetilde{E} := V^H EU$ and $\widetilde{F} := V^H FU$ can also be arbitrary $r \times r$ matrices such that $\|\widetilde{E}\|_2 \leq 1, \|\widetilde{F}\|_2 \leq \tau$ (for example, any $\widetilde{E}, \widetilde{F}$ are obtained by letting $E = V\widetilde{E}U^H$ and $F = V\widetilde{F}U^H$). Hence we see that the condition numbers $\kappa_i(A, B, \tau, \lambda_0)$ have the following characterization.

Lemma 2 *Under the assumptions in Lemma 1, suppose that $X_1 Y_1^H = U\Sigma V^H$ is an SVD where $\Sigma = \mathrm{diag}(\sigma_i) \in \mathbb{R}^{r \times r}$ ($\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$). Then*

$$\kappa_i(A, B, \tau, \lambda_0) = \sup_{\|\widetilde{E}\|_2 \leq 1, \|\widetilde{F}\|_2 \leq \tau} |\lambda_i(\Sigma(\widetilde{E} - \lambda_0 \widetilde{F}))|, \quad i = 1, \ldots, r. \quad (6)$$

In the sequel we use (6) to find closed-form expressions for $\kappa_i(A, B, \tau, \lambda_0)$.

3 Hermitian definite pairs

When (A, B) is a Hermitian definite pair, all the eigenvalues are always real and nondefective, and there exists a nonsingular matrix X such that [4]

$$X^H AX = \begin{bmatrix} \lambda_0 I_r & 0 \\ 0 & \Lambda_1 \end{bmatrix}, \quad X^H BX = I_n, \quad (7)$$

where Λ_1 is a diagonal matrix containing the eigenvalues not equal to λ_0. Hence the diagonals of Σ in (6) are the r positive singular values of the matrix $X_1 X_1^H$, which are equal to the eigenvalues of the matrix $X_1^H X_1$. Since (A, B) is a Hermitian definite pair it is natural to require that the perturbation matrices preserve the property, so (6) becomes the "structured" condition numbers $\kappa_i(A, B, \tau, \lambda_0; \mathbb{S})$, in which $\mathbb{S}$ indicates the Hermitian structure of the perturbations E, F (and hence $\widetilde{E}, \widetilde{F}$), expressed by

$$\kappa_i(A, B, \tau, \lambda_0; \mathbb{S}) \equiv \sup_{\substack{\|\widetilde{E}\|_2 \leq 1, \|\widetilde{F}\|_2 \leq \tau \\ \widetilde{E} = \widetilde{E}^H, \widetilde{F} = \widetilde{F}^H}} |\lambda_i(\Sigma(\widetilde{E} - \lambda_0 \widetilde{F}))|, \quad i = 1, \ldots, r. \quad (8)$$

The dependence of κ_i on the Hermitian structure $\mathbb{S}$ is made explicit because the structure affects κ_i in a nontrivial manner, as we shall see below.

Denoting

$$D = \Sigma^{1/2} = \mathrm{diag}(\sqrt{\sigma_1}, \ldots, \sqrt{\sigma_r}), \quad (9)$$

we see that the eigenvalues of $\Sigma(\widetilde{E} - \lambda_0 \widetilde{F})$ are equal to those of the Hermitian matrix $D(\widetilde{E} - \lambda_0 \widetilde{F})D$.

We further observe that $\widetilde{E} - \lambda_0 \widetilde{F}$ can represent an arbitrary Hermitian matrix H with $\|H\|_2 \leq 1 + \tau|\lambda_0|$, which can be done by letting $\widetilde{E} = H/\|H\|_2$ and $\widetilde{F} = -\tau\widetilde{E}|\lambda_0|/\lambda_0$. Conversely, it is easily seen that the class of Hermitian matrices H with $\|H\|_2 \leq 1 + \tau|\lambda_0|$ includes all the matrices expressed by $\widetilde{E} - \lambda_0 \widetilde{F}$. Together with the fact that the absolute values of the eigenvalues

of a Hermitian matrix are simply the singular values, we have yet another characterization of condition numbers in the Hermitian definite case:

$$\kappa_i(A, B, \tau, \lambda_0; \mathbb{S}) = (1 + \tau|\lambda_0|) \sup_{\substack{\|H\|_2 \leq 1 \\ H = H^H}} \sigma_i(DHD), \quad i = 1, \ldots, r. \qquad (10)$$

3.1 A closed-form expression of the condition numbers

Now we are ready to derive the expression of the condition numbers $\kappa_i(A, B, \tau, \lambda_0; \mathbb{S})$ using the singular values σ_i of $X_1 Y_1^H$.

Theorem 1 *Let (A, B) be a Hermitian definite pair that has the decomposition (7), so that λ_0 is a multiple eigenvalue of multiplicity r. Denote by σ_i the ith largest singular value of $X_1^H X_1$. Then the condition numbers of λ_0 under Hermitian perturbations $E = E^H, F = F^H$, as in (8), (10) are*

$$\kappa_i(A, B, \tau, \lambda_0; \mathbb{S}) = (1 + \tau|\lambda_0|) \min_{1 \leq j \leq i} \sqrt{\sigma_j \sigma_{i-j+1}}, \quad i = 1, \ldots, r. \qquad (11)$$

Note that $\sqrt{\sigma_j \sigma_{i-j+1}}$ is the geometric mean of σ_j and σ_{i-j+1}, the jth largest and smallest of $(\sigma_1, \sigma_2, \ldots, \sigma_i)$, which is the set of the i largest singular values of $X_1^H X_1$.

Proof In view of (10), to prove the theorem it suffices to prove that for D as in (9) and for any Hermitian H such that $\|H\|_2 = 1$, $\sigma_i(DHD)$ is bounded above by $\min_j \sqrt{\sigma_j \sigma_{i-j+1}}$ for $i = 1, \ldots, r$, and that this bound is attainable.

First, proving attainability is simply done by considering the case where H is zero except for its $i \times i$ leading principal submatrix, which has 1s on the antidiagonal and 0 elsewhere, that is, $H_{j,k} = 1$ if $j + k = i + 1$ and otherwise $H_{j,k} = 0$. This choice of H makes the $i \times i$ leading principal submatrix of DHD also an anti-diagonal matrix, whose jth antidiagonal is $\sqrt{\sigma_j \sigma_{i-j+1}}$. Hence we have $\sigma_i(DHD) = \min_j \sqrt{\sigma_j \sigma_{i-j+1}}$.

Our remaining task is to prove that $\min_j \sqrt{\sigma_j \sigma_{i-j+1}}$ is an upper bound of $\sigma_i(DHD)$ for any Hermitian H with $\|H\|_2 \leq 1$. Using the max-min characterization of singular values [11, p. 68], we have

$$\sigma_i(DHD) = \max_{Q^H Q = I_i} \min_{\|v\|_2 = 1} \|DHDQv\|_2,$$

so it suffices to show that for any $Q \in \mathbb{C}^{r \times i}$ with orthonormal columns, there exists a unit vector v such that $\|DHDQv\|_2 \leq \min_j \sqrt{\sigma_j \sigma_{i-j+1}}$.

To prove this, let $j_0 = \mathrm{argmin}_{j \leq (i+1)/2} \sqrt{\sigma_j \sigma_{i-j+1}}$. Since for any Q we have $\mathrm{rank}(Q(1:i-j_0,:)) \leq i - j_0$, there are at least j_0 linearly independent vectors in $\mathbb{C}^{i \times 1}$ that are orthogonal to the rows of $Q(1:i-j_0,:)$. Therefore there must exist $P \in \mathbb{C}^{i \times j_0}$ with orthonormal columns such that the first $i - j_0$ rows of the $r \times j_0$ matrix QP are all zeros. For such P, we have

$$\|DQP\|_2 = \|\mathrm{diag}(0, \ldots, 0, \sqrt{\sigma_{i-j_0+1}}, \ldots, \sqrt{\sigma_r})QP\|_2 \leq \sqrt{\sigma_{i-j_0+1}}.$$

Furthermore, since denoting by $W = H(1: j_0 - 1, :)DQP$ the first $j_0 - 1$ rows of $HDQP$ we have $\text{rank}(W) \leq j_0 - 1$, so there must exist a unit vector $w \in \mathbb{C}^{j_0 \times 1}$ that is orthogonal to W, so that the first $j_0 - 1$ rows of $HDQPw$ are all zeros. We easily see that for such w we have $\|DHDQPw\|_2 \leq \sqrt{\sigma_{j_0} \sigma_{i-j_0+1}}$. Therefore we have shown that for any $Q \in \mathbb{C}^{k \times i}$ with orthonormal columns there exists a unit vector $v_0 = Pw$ such that

$$\min_{\|v\|_2=1} \|DHDQv\|_2 \leq \|DHDQv_0\|_2 \leq \sqrt{\sigma_{j_0} \sigma_{i-j_0+1}} = \min_j \sqrt{\sigma_i \sigma_{i-j+1}}.$$

$\square$

Three remarks are in order.

– When $B \neq I_n$, σ_i for $i = 1, \ldots, r$ generally take different values, so (11) shows that a multiple eigenvalue in generalized eigenvalue problems actually take multiple condition numbers that are different from each other, which is our main result. Note that the ratio among the condition numbers is bounded by $\kappa_1(A, B, \tau, \lambda_0; \mathbb{S})/\kappa_r(A, B, \tau, \lambda_0; \mathbb{S}) \leq \sigma_1/\sigma_r$. Now since $\sigma_{\min}(B^{-1}) \leq \sigma_r \leq \sigma_1 \leq \sigma_{\max}(B^{-1})$, we have $\sigma_1/\sigma_r \leq \sigma_{\max}(B)/\sigma_{\min}(B) = \kappa_2(B)$, the standard 2-norm condition number of B. It follows that if B is well-conditioned then a multiple eigenvalue of a Hermitian definite pair must have similar condition numbers.

– The matrix H that attains the supremum in (10) for each i is highly indefinite, suggesting that the supremums cannot be attained when the perturbation is definite. In fact, if $H = \widetilde{E} - \lambda_0 \widetilde{F}$ is (positive) definite, then DHD is also positive definite, so the singular values are equal to the eigenvalues, and noting that $D^2\|H\|_2 - DHD = D(\|H\|_2 I - H)D$ is positive semidefinite we conclude that $\sigma_i(DHD) \leq \sigma_i(D^2\|H\|_2) = \sigma_i$. Hence if we confine to perturbation where H is definite, the condition numbers become $(1 + \tau|\lambda_0|)\sigma_i$.

– The above arguments show that the difference among condition numbers of a multiple eigenvalue is due to the difference among the r singular values of $X_1 Y_1^H$, the outer product of the left and right eigenvectors corresponding to λ_0. $\sigma_i(X_1 Y_1^H)$ are all 1 in the standard Hermitian case because $X_1 = Y_1$ has orthonormal columns. In the standard non-Hermitian case $X_1 \neq Y_1$ and neither has orthonormal columns, so the $\sigma_i(X_1 Y_1^H)$ generally take r different values. In the generalized Hermitian case we have $X_1 = Y_1$ but X_1 has B-orthonormal columns that are not orthonormal, so $\sigma_i(X_1 X_1^H)$ again can take r different values. Note that if one uses the B-based inner product the difference in condition numbers disappears (recall the remark in the introduction).

4 Non-Hermitian pairs

Here we consider the case where (A, B) is a general non-Hermitian pair. In view of (6), our task is to bound $|\lambda_i(X\Sigma)|$ for an arbitrary square matrix X such that $\|X\|_2 \leq 1$. This is in fact the exact same problem addressed in

[16, Thm. 3.1]. Hence the analysis there can be directly applied to yield the following result.

Theorem 2 *Let (A, B) be a non-Hermitian pair that satisfies (3), and let $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ be the positive singular values of the matrix $X_1 Y_1^H$. Then, $\kappa_i(A, B, \tau, \lambda_0)$ under non-Hermitian perturbations can be expressed as*

$$\kappa_i(A, B, \tau, \lambda_0) = (1 + \tau|\lambda_0|)\left(\prod_{j=1}^{i} \sigma_j\right)^{1/i}, \quad i = 1, \ldots, r. \tag{12}$$

4.1 Structured perturbations

It is instructive to revisit the Hermitian definite case, but now allowing for non-Hermitian perturbations, that is, E, F are general matrices whose norms are bounded by 1. In this case, the condition numbers $\kappa_i(A, B, \tau, \lambda_0)$ have the characterization (6) (instead of (10)), so they have the expression (12), the same as that for a non-Hermitian pair.

As might be expected, the condition number under Hermitian perturbation (11) is never larger than that under non-Hermitian perturbations (12):

$$\frac{\kappa_i(A, B, \tau, \lambda_0; \mathbb{S})}{1 + \tau|\lambda_0|} = \left(\min_{1 \leq j \leq i}(\sigma_j \sigma_{i-j+1})^i\right)^{1/2i} \leq \left(\prod_{j=1}^{i}(\sigma_j \sigma_{i-j+1})\right)^{1/2i}$$

$$= \left(\prod_{j=1}^{i} \sigma_j^2\right)^{1/2i} = \left(\prod_{j=1}^{i} \sigma_j\right)^{1/i} = \frac{\kappa_i(A, B, \tau, \lambda_0)}{1 + \tau|\lambda_0|}.$$

The above arguments imply that under general non-Hermitian perturbation the Hermitian structure of A and B plays no direct role in the condition numbers. In particular, for a Hermitian definite pair (A, B) and a non-Hermitian pair $(\widetilde{A}, \widetilde{B})$ that both have a multiple eigenvalue λ_0 of multiplicity r with normalized eigenvectors X_1, Y_1 and $\widetilde{X}_1, \widetilde{Y}_1$ respectively, if $\sigma_i(X_1 Y_1^H) = \sigma_i(\widetilde{X}_1 \widetilde{Y}_1^H)$ for $i = 1, \ldots, r$, then under a general non-Hermitian perturbation we have $\kappa_i(A, B, \tau, \lambda_0) = \kappa_i(\widetilde{A}, \widetilde{B}, \tau, \lambda_0) = (1 + \tau|\lambda_0|)(\prod_{j=1}^{i} \sigma_j)^{1/i}$ for $i = 1, \ldots, r$. Therefore, for a Hermitian definite pair the structure of the perturbation matrices plays an important role in the sensitivity of a multiple eigenvalue. We note that the standard Hermitian case with $B \equiv I$ is an exception, in which the condition numbers are always all 1 whether or not the perturbation matrices are Hermitian.

This point of view, to focus on the effect of the structure of the perturbation, was investigated extensively in [6], in which (Theorem 4.5) it is shown that (among other structures they consider) the Hermitian structure of the perturbation matrices does not have any effect on the Hölder condition number.

At first sight this seems to contradict our results, which show that the Hermitian structure of the perturbation matrices does affect the condition

numbers of the multiple eigenvalue λ_0. The explanation is that [6] treats only the Hölder condition number, which is equivalent to $\kappa_1(A, B, \tau, \lambda_0)$ in the nondefective case. Here we are identifying individual condition numbers of each of the r eigenvalues. In fact, we can see that for $i = 1$, κ_i in (11) and (12) are the same, both equal to $(1 + \tau|\lambda_0|)\sigma_1$. We can easily see that they are equal also for $i = 2$. The difference between (11) and (12) starts to take effect only for $i \geq 3$, so λ_0's multiplicity r must be at least 3. In particular, for a simple eigenvalue the Hermitian structure of the perturbation has no effect on the condition number, which is a trivial consequence of the results in [6], and also pointed out in [5].

4.2 Examples

Here we present two simple examples to illustrate the above results and observations, and to reveal the behavior of a multiple eigenvalue under random perturbations.

Example 1 For the Hermitian definite pair $\left[\begin{smallmatrix} 2000 & 0 \\ 0 & 2 \end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix} 1000 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ presented in [13, p. 300], we have $\kappa_1(A, B, \tau, \lambda_0; \mathbb{S}) = \kappa_1(A, B, \tau, \lambda_0) = 3$ and $\kappa_2(A, B, \tau, \lambda_0; \mathbb{S}) = \kappa_2(A, B, \tau, \lambda_0) = 3/\sqrt{1000}$, which shows that the multiple eigenvalue $\lambda_0 = 2$ has different sensitivities. Note that in this case the structure of the perturbation has no effect on the condition numbers, because the multiplicity of λ_0 is $r < 3$.

Example 2 We consider a 4×4 Hermitian definite pair (A, B) expressed by

$$A = W^H \Lambda W, \quad B = W^H W,$$

where $\Lambda = \mathrm{diag}(1, 1, 1, 2)$ and $W = \mathrm{diag}(1, 2, 100, 1)$, so the eigenvalues of (A, B) are $1, 1, 1, 2$. Since X that diagonalizes A, B (as in (7)) is $X = W^{-1} = \mathrm{diag}(1, 0.5, 0.01, 1)$ and X_1 is its first three columns, the singular values of $X_1 X_1^H$ are $\sigma_1 = 1^2, \sigma_2 = 0.5^2, \sigma_3 = 0.01^2$ (where in this example we let $\tau = 1$), hence by (11) it follows that $\kappa_1(A, B, 1, 1; \mathbb{S}) = 2, \kappa_2(A, B, 1, 1; \mathbb{S}) = 1$ and $\kappa_3(A, B, \tau, 1; \mathbb{S}) = 0.02$. Using MATLAB version 7.10 we generated 10^6 sets of random Hermitian perturbation matrices E and F such that $||E||_2, ||F||_2 \leq 1$ (whose elements were generated using the MATLAB command `randn`, then normalizing $E \leftarrow \zeta E/||E||_2$ where $\zeta = $ `rand` ≤ 1 is another positive random number), and examined the behavior of the three eigenvalues of the pair $(A + \epsilon E, B + \epsilon F)$ that are closest to $\lambda_0 = 1$, where we let $\epsilon = 10^{-5}$. Specifically, recalling the discussion in the introduction, we denote by $1 + \Delta_i(\epsilon, E, F)$ for $i = 1, 2, 3$ the three eigenvalues of $(A + \epsilon E, B + \epsilon F)$ that are closest to 1 such that $|\Delta_1(\epsilon, E, F)| \geq |\Delta_2(\epsilon, E, F)| \geq |\Delta_3(\epsilon, E, F)|$, and examine how large $|\Delta_i(\epsilon, E, F)|/\epsilon$ can be for $i = 1, 2, 3$.

We also experimented with non-Hermitian perturbations, in which case we let E, F be random non-Hermitian matrices with $||E||_2, ||F||_2 \leq 1$. In this case the condition numbers (12) are $\kappa_1(A, B, 1, 1) = 2, \kappa_2(A, B, 1, 1) = $

$2(1 \cdot 0.5^2)^{1/2} = 1$ and $\kappa_3(A, B, 1, 1) = 2(1 \cdot 0.5^2 \cdot 0.01^2)^{1/3} \simeq 0.058$, in which we confirm that the first two are the same as in the above Hermitian case.

Lastly, in order to see how the Hermitian property of the matrices plays a role in the eigenvalue perturbation behaviors, we also tested with a non-Hermitian pair (A, B) that has the same eigenvalues and σ_i (of $X_1 Y_1^H$) as the above Hermitian pair. We formed such a pair (A, B) by defining $A = Y^{-H} \Lambda X^{-1}$ and $B = Y^{-H} X^{-1}$, where $\Lambda = \mathrm{diag}(1, 1, 1, 2)$, Y_1^H (the first 3 rows of Y) is set to $Z \Sigma V^H$ and X_1 (the first 3 columns of X) is set to $U Z^{-1}$, where U and V are randomly generated matrices with orthonormal columns, $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3) = (1^2, 0.5^2, 0.01^2)$ and Z is an arbitrary nonsingular matrix[1]. Elements of the last row of Y and the last column of X were taken as random numbers. Since we have $X_1 Y_1^H = U \Sigma V^H$, we have $\kappa_1(A, B, 1, 1) = 2, \kappa_2(A, B, 1, 1) = 1$ and $\kappa_3(A, B, 1, 1) = 0.058$, the same condition numbers as the above second case with non-Hermitian perturbation, as was intended. The perturbations E and F are taken as arbitrary non-Hermitian matrices.

In summary, we tested three different situations, all of which have the same $\sigma_i(X_1 Y_1^H)$: (i) Both (A, B) and (E, F) are Hermitian (shown as "Herm + Herm" in Table 2), (ii) (A, B) is Hermitian but (E, F) is non-Hermitian ("Herm + NonHerm"), and (iii) Both (A, B) and (E, F) are non-Hermitian ("NonHerm + NonHerm").

The results are summarized in Table 2 below, which shows the average and maximum (shown as avg. and max respectively) values of $|\Delta_i(\epsilon, E, F)|/\epsilon$ among the 10^6 runs with randomly generated E and F, along with the condition numbers κ_i (which is a first order upper bound for $|\Delta_i(\epsilon, E, F)|/\epsilon$) for $i = 1, 2, 3$.

Table 2 Average and maximum perturbation $|\Delta_i(\epsilon, E, F)|/\epsilon$ of 10^6 runs for $i = 1, 2, 3$.

i	Herm + Herm			Herm + NonHerm			NonHerm + NonHerm		
	avg.	max	κ_i	avg.	max	κ_i	avg.	max	κ_i
1	0.579	1.98	2.0	0.41	1.86	2.0	0.42	1.90	2.0
2	0.141	0.84	1.0	0.136	0.76	1.0	0.137	0.76	1.0
3	0.00018	0.012	0.02	0.00021	0.027	0.058	0.00021	0.027	0.058

We make the following observations.

– We confirm that κ_i is an upper bound of $\max |\Delta_i(\epsilon, E, F)|/\epsilon$ for all i in all three cases (which is necessarily true in the limit $\epsilon \to 0$). Interestingly, for $i = 1$ the bound κ_i is nearly attained while for $i = 2, 3$, $\max |\Delta_i(\epsilon, E, F)|/\epsilon$ is noticeably smaller than κ_i, which suggests that for larger i it becomes more and more rare that the largest possible perturbation is attained.

– Reflecting the fact that κ_i are the same for all the three cases for $i = 1$ and 2, we can see that $\max |\Delta_i(\epsilon, E, F)|/\epsilon$ are similar in all three cases, so two eigenvalues have similar maximum sensitivities regardless of whether

---

[1] Note that the choice of Z does not affect the condition numbers $\kappa_i(A, B, 1, 1)$.

A, B, E, F are Hermitian or not. On the contrary, $\max |\Delta_i(\epsilon, E, F)|/\epsilon$ for $i = 3$ show the somewhat different sensitivities of the third eigenvalue depending on the structure of E, F.

– The behavior of the multiple eigenvalue is remarkably similar for the latter two cases, not only in terms of $\max |\Delta_i(\epsilon, E, F)|/\epsilon$ but also $\mathrm{avg.}|\Delta_i(\epsilon, E, F)|/\epsilon$. This reflects the fact that the first order expansions of λ_0 are the same for the two cases, so that the local behavior of an eigenvalue is determined solely by the singular values of $X_1 Y_1^H$, which does not depend explicitly on whether A and B are Hermitian.

– Comparing $\mathrm{avg.}|\Delta_i(\epsilon, E, F)|/\epsilon$ with $\max |\Delta_i(\epsilon, E, F)|/\epsilon$, we see that the former is much smaller than the latter for larger i. For $i = 1$ the difference seems less significant.

A precise explanation for the last two observations, which necessarily involves statistical analysis, is an open problem: our discussions deal only with the maximum attainable perturbation $\max |\Delta_i(\epsilon, E, F)|/\epsilon$, not with $\mathrm{avg.}|\Delta_i(\epsilon, E, F)|/\epsilon$.

It is important to note that the above experiments only exhibit the behavior of a multiple eigenvalue under random perturbations (artificially constructed from a certain type of random matrices), which do not necessarily reflect the typical backward errors observed in practical numerical methods.

4.3 Defective and infinite cases

So far we have treated only the case where λ_0 is a finite and nondefective multiple eigenvalue. Here we briefly consider the cases where λ_0 is infinite and/or defective.

The case $\lambda_0 = \infty$ can be treated as in [2,6] simply by considering the multiple zero eigenvalue of the pair (B, A), for which the exact same discussion as above is valid.

When λ_0 is defective, Lidskii's perturbation theory [2,6,8] shows that the leading term in λ_0's perturbation expansion is not linear in ϵ. Specifically, if λ_0 is an eigenvalue of (A, B) of multiplicity $n_1 r$ belonging to a Jordan block of dimension n_1 repeated r times then there are $n_1 r$ eigenvalues of $(A + \epsilon E, B + \epsilon F)$ admitting the expansion

$$\widehat{\lambda}_{i,\ell} = \lambda_0 + \left(\lambda_i(Y_1^H(E - \lambda_0 F)X_1)\right)^{1/n_1} \epsilon^{1/n_1} + o(\epsilon^{1/n_1}) \qquad (13)$$

for $i = 1, 2, \ldots, r$ and $\ell = 1, 2, \ldots, n_1$. Here $Y_1^H \in \mathbb{C}^{r \times n}$ and $X_1 \in \mathbb{C}^{n \times r}$ represent the linearly independent left and right eigenvectors of (A, B) corresponding to λ_0, and the value $\left(\lambda_i(Y_1^H(E - \lambda_0 F)X_1)\right)^{1/n_1}$ represents all the n_1 distinct n_1th roots.

In view of the exponent in (13), we define the condition numbers of λ_0 with the exponent $1/n_1$ by

$$\kappa_{i,n_1}(A, B, \tau, \lambda_0) = \lim_{\epsilon \to 0} \sup_{\|E\|_2=1, \|F\|_2=\tau} \frac{|\Delta_{i,\ell}(\epsilon, E, F)|}{|\epsilon|^{1/n_1}} \qquad (14)$$

for $i = 1, \ldots, r$ and $\ell = 1, \ldots, n_1$, where $\Delta_{i,\ell}(\epsilon, E, F) = \widehat{\lambda}_{i,\ell} - \lambda_0$ such that $|\Delta_{1,\ell}(\epsilon, E, F)| \geq \cdots \geq |\Delta_{r,\ell}(\epsilon, E, F)|$, $\lim_{\epsilon \to 0} \Delta_{i,\ell}(\epsilon, E, F) = 0$ and $\lim_{\epsilon \to 0} |\Delta_{i,\ell}(\epsilon, E, F)|/|\Delta_{i,\widetilde{\ell}}(\epsilon, E, F)| = 1$ for any $i \in \{1, \ldots, r\}$ and $\ell, \widetilde{\ell} \in \{1, \ldots, n_1\}$.

By (13) and (14) we have $\kappa_{i,n_1}(A, B, \tau, \lambda_0) = \sup_{\|E\|_2 \leq 1, \|F\|_2 \leq \tau} |\lambda_i(Y_1^H(E - \lambda_0 F)X_1)|^{1/n_1}$. We observe that although the leading exponent is different from that in (4), the sensitivities of the multiple eigenvalue are still governed by $|\lambda_i(Y_1^H(E - \lambda_0 F)X_1)|$ for $i = 1, \ldots, r$, for which we gave a bound in the above discussions.

Hence as in Theorem 2, by (13) and (14) we have

$$\kappa_{i,n_1}(A, B, \tau, \lambda_0) = \left((1 + \tau|\lambda_0|) \big(\prod_{j=1}^{i} \sigma_j \big)^{1/i} \right)^{1/n_1}.$$

By (13), we must have $|\widehat{\lambda}_{i,\ell} - \lambda_0|/\epsilon^{1/n_1} \leq \kappa_{i,n_1}(A, \tau, B, \lambda_0)$ for $i = 1, \ldots, r$ in the limit $\epsilon \to 0$ for any E and F. Note that $\kappa_{i,n_1}(A, B, \tau, \lambda_0)$ does not depend explicitly on ℓ (which is why we do not write $\kappa_{i,n_1,\ell}(A, B, \tau, \lambda_0)$), which means the n_1 condition numbers of the eigenvalues belonging to a certain Jordan block are the same. We observe that in the defective case $n_1 \geq 2$, the exponent $1/n_1$ makes the difference among the condition numbers less significant than in the nondefective case, see the example below for an illustration.

4.3.1 Example

To examine the behavior of a defective multiple eigenvalue, we generate a 7×7 pair (A, B) defined by

$$A = Y^{-H} \begin{bmatrix} J & & & \\ & J & & \\ & & J & \\ & & & 2 \end{bmatrix} X^{-1}, \quad \text{and} \quad B = Y^{-H}X^{-1}, \tag{15}$$

where $J = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is a 2×2 Jordan block. (A, B) has a multiple eigenvalue $\lambda_0 = 1$ of multiplicity six and a simple eigenvalue 2. $Y_1^H \equiv [Y(:, 2)\ Y(:, 4)\ Y(:, 6)]^H = Z\Sigma V^H$ and $X_1 \equiv [X(:, 1)\ X(:, 3)\ X(:, 5)] = UZ^{-1}$ are the left and right eigenvectors corresponding to λ_0, where U and V are random matrices with orthonormal columns and Z is an arbitrary nonsingular matrix. The other rows of Y^H and columns of X do not affect the condition numbers of λ_0, so we let them take random values. We let $\Sigma = \text{diag}(1^2, 0.5^2, 0.01^2)$, so that $\sigma_i(X_1 Y_1^H) = \{1^2, 0.5^2, 0.01^2\}$ are the same values as in the non-Hermitian case of the second example in section 4.2.

Recall from (13) that perturbations in (A, B) generally split λ_0 into $n_1 r$ perturbed eigenvalues $\widehat{\lambda}_{i,\ell}$ for $i = 1, \ldots, r$ and $\ell = 1, \ldots, n_1$. (13) also shows that for a fixed i, $|\widehat{\lambda}_{i,\ell} - \lambda_0|$ differ by at most $o(\epsilon^{1/n_1})$ for varying ℓ. For the matrix pair (15) we have $r = 3$ and $n_1 = 2$, so we separate the six eigenvalues

$\widehat{\lambda}_{i,\ell}$ into three groups according to the value of i, so that the two eigenvalues of the ith group have sensitivity governed by $\left|\lambda_i(Y_1^H(E - \lambda_0 F)X_1)\right|^{1/n_1}$.

With $\tau = 1$, the condition numbers $\kappa_{i,2}(A, B, 1, 1)$ for the ith group for $i = 1, 2, 3$ are $\kappa_{1,2}(A, B, 1, 1) = (2\cdot 1)^{1/2} = \sqrt{2}$, $\kappa_{2,2}(A, B, 1, 1) = \left(2 \cdot (1 \cdot 0.5^2)^{1/2}\right)^{1/2} = 1$ and $\kappa_{3,2}(A, B, 1, 1) = \left(2 \cdot (1 \cdot 0.5^2 \cdot 0.01^2)^{1/3}\right)^{1/2} \simeq 0.24$. Comparing these with $\kappa_i(A, B, 1, 1)$ in the example in section 4.2 we see that that although $\sigma_i(X_1 Y_1^H)$ take the same values, the relative difference among the condition numbers $\kappa_{1,n_1}/\kappa_{r,n_1}$ is smaller than $\kappa_{1,1}/\kappa_{r,1}$ for $n_1 \geq 2$, due to the exponent $1/2$.

Recalling that we must have $|\widehat{\lambda}_{i,\ell} - 1|/\epsilon^{1/2} \leq \kappa_{i,2}(A, B, 1, 1)$ for small ϵ, here we examine how large $|\widehat{\lambda}_{i,\ell} - 1|/\epsilon^{1/2}$ becomes for $i = 1, 2, 3$. To do this, of the six eigenvalues of $(A + \epsilon E, B + \epsilon F)$ close to λ_0, we check the perturbation of the most perturbed, third perturbed, and the fifth perturbed ones.

In the experiment we let E, F be random non-Hermitian matrices with $||E||_2, ||F||_2 \leq 1$, let $\epsilon = 10^{-6}$ and tested with 10^6 pairs. In Table 3 we report the average and maximum values of $|\widehat{\lambda}_{i,\ell} - 1|/\epsilon^{1/2}$ for $i = 1, 2, 3$.

Table 3 Defective matrix pair (15) with three 2×2 Jordan blocks, average and maximum perturbation $|\widehat{\lambda}_{i,\ell} - 1|/\epsilon^{1/2}$ of 10^6 runs for $i = 1, 2, 3$.

i	avg.	max	$\kappa_{i,2}$
1	0.511	1.21	1.41
2	0.282	0.733	1
3	0.0089	0.138	0.24

As in the experiments in section 4.2 for nondefective multiple eigenvalues, we see that a defective multiple eigenvalue also exhibits different sensitivities under perturbation. We also tested with Hermitian perturbations $E = E^H$ and $F = F^H$, and obtained nearly the same results as in Table 3, suggesting that the structure of the perturbation does not play a role here.

In all our experiments we had $|\widehat{\lambda}_{i,1} - \widehat{\lambda}_{i,2}|/\epsilon^{1/2} < 0.04$ for $i = 1, 2, 3$, which matches the theoretical result indicated by (13) that for a given i, $|\widehat{\lambda}_{i,\ell} - 1|$ differ only by $o(\epsilon^{1/n_1})$ for varying ℓ.

Finally, comparing Table 3 and the third case of Table 2 we see that the difference among each column is less pronounced in Table 3. This suggests that the difference in sensitivities among the multiple eigenvalue λ_0 is smaller in the defective case $n_1 \geq 2$, reflecting the last remark before this example.

References

1. Barlow, J., Demmel, J.: Computing accurate eigensystems of scaled diagonally dominant matrices. SIAM J. Numer. Anal. **27**(3), 762–791 (1990)
2. De Teran, F., Dopico, F.M., Moro, J.: First order spectral perturbation theory of square singular matrix pencils. Linear Algebra Appl. **429**(2-3), 548–576 (2008)
3. Demmel, J., Veselic, K.: Jacobi's method is more accurate than QR. SIAM J. Matrix Anal. Appl. **13**(4), 1204–1245 (1992)
4. Golub, G.H., Van Loan, C.F.: Matrix Computations, third edn. The Johns Hopkins University Press (1996)
5. Higham, D.J., Higham, N.J.: Structured backward error and condition of generalized eigenvalue problems. SIAM J. Matrix Anal. Appl. **20**(2), 493–512 (1998)
6. Kressner, D., Jose Pelaez, M., Moro, J.: Structured Hölder condition numbers for multiple eigenvalues. SIAM J. Matrix Anal. Appl. **31**(1), 175–201 (2009)
7. Li, R.C., Nakatsukasa, Y., Truhar, N., Xu, S.: Perturbation of partitioned Hermitian definite generalized eigenvalue problems. SIAM J. Matrix Anal. Appl. **32**(2), 642–663 (2011)
8. Moro, J., Burke, J.V., Overton, M.L.: On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. SIAM J. Matrix Anal. Appl. **18**(4), 793–817 (1997)
9. Nakatsukasa, Y.: Perturbation behavior of a multiple eigenvalue in generalized Hermitian eigenvalue problems. BIT **50**(1), 109–121 (2010)
10. Stewart, G.W.: Error and perturbation bounds for subspaces associated with certain eigenvalue problems. SIAM Rev. **15**(4), pp. 727–764 (1973)
11. Stewart, G.W.: Matrix Algorithms Volume I: Basic Decompositions. SIAM (1998)
12. Stewart, G.W.: Matrix Algorithms Volume II: Eigensystems. SIAM (2001)
13. Stewart, G.W., Sun, J.G.: Matrix Perturbation Theory. Academic Press (1990)
14. Stewart, G.W., Zhang, G.: Eigenvalues of graded matrices and the condition numbers of a multiple eigenvalue. Numer. Math. **58**(7), 703–712 (1991)
15. Sun, J.G.: On condition numbers of a nondefective multiple eigenvalue. Numer. Math. **61**(2), 265–275 (1992)
16. Sun, J.G.: On worst-case condition numbers of a nondefective multiple eigenvalue. Numer. Math. **69**(3), 373–382 (1995)
17. Veselic, K., Slapnicar, I.: Floating-point perturbations of Hermitian matrices. Linear Algebra Appl. **195**, 81 – 116 (1993)