Sensitivity and Computation of a Defective Eigenvalue

Zhonggang Zeng*

April 9, 2015

Abstract

A multiple and defective eigenvalue is well documented to be hypersensitive to data perturbations and round-off errors, making it a formidable challenge in numerical computation particularly when entries of the matrix are not known exactly. This paper establishes a finitely bounded sensitivity of a defective eigenvalue with respect to perturbations that preserve the geometric multiplicity and the smallest Jordan block size. Based on this perturbation theory, numerical computation of a defective eigenvalue is regularized as a well-posed least squares problem so that it can be accurately carried out using floating point arithmetic even if the matrix is perturbed.

1 Introduction

Computing matrix eigenvalues is one of the fundamental problems in theoretical and numerical linear algebra. The highly celebrated Francis QR algorithm effectively solves the problem for simple and non-defective eigenvalues in practical applications. However, it is well documented that multiple and defective eigenvalues are hypersensitive to both data perturbations on the matrix and the inevitable round-off. For an eigenvalue of a matrix A associated with the largest Jordan block size $l \times l$ while A is perturbed by ΔA, the error bound [2, p. 58][3, 13] on the eigenvalue deviation is a constant multiple of $\|\Delta A\|_2^{\frac{1}{l}}$, implying that the accuracy of the computed eigenvalue in the number of correct digits is only a fraction $\frac{1}{l}$ of the accuracy of the matrix data. As a result, numerical computation of defective eigenvalues from empirical matrix data remains a formidable challenge.

On the other hand, it has long been known that a defective eigenvalue disperses into a cluster of simple eigenvalues when the matrix is under arbitray perturbations but the mean of the cluster is not hypersensitive [10, 17]. In his seminal technical report[9], Kahan proved that the sensitivity of an m-fold eigenvalue is actually bounded by $\frac{1}{m}\|P\|_2$ where P is the spectral projector associated with the eigenvalue as long as the perturbation is

*Department of Mathematics, Northeastern Illinois University, Chicago, Il 60625. *Email:* zzeng@neiu.edu

constrained so that the multiplicity is preserved. The same proof and the same sensitivity also apply to the mean of the eigenvalue cluster emanating from the m-fold eigenvalue with respect to arbitrary perturbations. Indeed, using cluster means as approximations to defective eigenvalues has been extensively applied to numerical computation of Jordan Canonical Forms and staircase forms, provided that the clusters can be sorted out from the spectrum. This approach includes works of Ruhe [16], Sdridhar and Jordan [20], and culminated in Golub and Wilkinson's review [6] as well as Kågström and Ruhe's JNF [7, 8]. Theoretical issues have been analyzed in, e.g. works of Demmel [4, 5] and Wilkinson [22, 23]. Perturbations on eigenvalue clusters are also studied as pseudospectra of matrices in works of Trefethon and Ebree [21] as well as Rump [18, 19].

In this paper we elaborate a different measurement of the sensitivity of a defective eigenvalue with respect to perturbations constrained to preserve the geometric multiplicity and the smallest Jordan block size. We prove that such a sensitivity is also finitely bounded even if the multiplicity is not preserved, and it is large only if either the geometric multiplicity or the smallest Jordan block size can be increased by a small perturbation on the matrix. This sensitivity can be small even if the spectral projector norm is large, or vice versa.

In general, perturbations are expected to be arbitrary without preserving either the multiplicity or what we refer to as the multiplicity support. We prove that a certain type of pseudo-eigenvalue uniquely exists, is Lipschitz continuous, is backward accurate and approximates the defective eigenvalue with a forward accuracy in the same order of the data accuracy. Namely, we prove that finding a defective eigenvalue via computing a pseudo-eigenvalue is a well-posed problem. Based on this analysis, we develop an iterative algorithm that is capable of accurate computation of defective eigenvalues using floating point arithmetic from empirical matrix data even if the spectral projector norm is large and thus the cluster mean is inaccurate.

2 Notation

The space of dimension n column vectors is $\mathbb{C}^n$ and the vector space of $m \times n$ matrices is $\mathbb{C}^{m\times n}$ as usual. Matrices are denoted by upper case letters A, X, and G, etc, with O representing a zero matrix whose dimensions can be derived from the context. Boldface lower case letters such as $\mathbf{x}$ and $\mathbf{y}$ represent column vectors. Particularly, the zero vector in $\mathbb{C}^n$ is denoted by $\mathbf{0}_n$ or simply $\mathbf{0}$ if the dimension is clear. The conjugate transpose of a matrix or vector $(\cdot)$ is denoted by $(\cdot)^{\mathsf{H}}$, and the pseudo-inverse of a matrix $(\cdot)$ is $(\cdot)^{\dagger}$. The submatrix formed by entries in rows $i_1, \ldots, i_2$ and columns $j_1, \ldots, j_2$ of a matrix A is denoted by $A_{i_1:i_2,j_1:j_2}$. The kernel and range of a matrix $(\cdot)$ are denoted by $\mathit{Ker}(\cdot)$ and $\mathit{Ran}(\cdot)$ respectively. The notation $\mathit{eig}(\cdot)$ represents the spectrum of a matrix $(\cdot)$.

We also consider vectors in vector spaces such as $\mathbb{C} \times \mathbb{C}^{m\times k}$ throughout this paper. In such cases, the vector 2-norm is the square root of the sum of squares of all components. For instance, a vector $(\lambda, X) \in \mathbb{C} \times \mathbb{C}^{n\times k}$ can be arranged as a column vector $\mathbf{u}$ in $\mathbb{C}^{n\,k+1}$ and $\|(\lambda, X)\|_2 = \|\mathbf{u}\|_2$ regardless of the ordering. A zero vector in such a vector space is also denoted by $\mathbf{0}$.

Let λ_* be an eigenvalue of a matrix $A \in \mathbb{C}^{n\times n}$. Its algebraic multiplicity can be partitioned into a non-increasing sequence of nonnegative integers called the *Segre characteristic* $\{l_1 \geq l_2 \geq \cdots\}$ that are the sizes of elementary Jordan blocks. In other words, there is a matrix $X_* \in \mathbb{C}^{n\times m}$ of full column rank such that

$$A\,X_* \;=\; X_* \begin{bmatrix} J_{l_1}(\lambda_*) & & \\ & J_{l_2}(\lambda_*) & \\ & & \ddots \end{bmatrix}$$

where

$$J_k(\lambda_*) \;=\; \begin{bmatrix} \lambda_* & 1 & & \\ & \lambda_* & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_* \end{bmatrix} \;\in\; \mathbb{C}^{k\times k}$$

denotes the $k \times k$ elementary Jordan block associated with an eigenvalue λ_*. For convenience, a Segre characteristic is infinite in formality and the number of nonzero entries is the geometric multiplicity. The last nonzero component of a Segre characteristic, namely the size of the smallest elementary Jordan block associated with λ_*, is of particular importance in our perturbation analysis and we shall call it the *Segre characteristic anchor*.

For instance, if λ_* is an eigenvalue of A associated with elementary Jordan blocks $J_5(\lambda_*), J_5(\lambda_*), J_4(\lambda_*), J_4(\lambda_*)$ and $J_3(\lambda_*)$, its Segre characteristic is $\{5, 5, 4, 4, 3, 0, \ldots\}$ with an anchor 3. The geometric multiplicity is 5 since the Segre characteristic anchor is the fifth component. A Segre characteristic along with its conjugate that is called the Weyr characteristic can be illustrated by a Ferrer's diagram in Fig. 1, where the geometric multiplicity and the Segre characteristic anchor represent the dimensions of the base rectangle occupied by the equal leading entries of the Weyr characteristic.

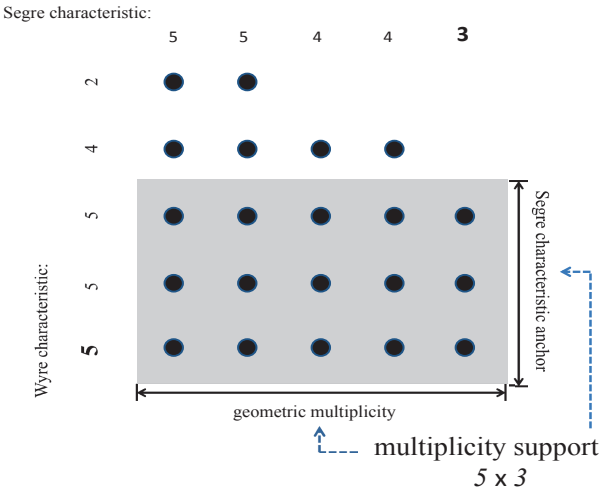

Figure 1: Illustration of the multiplicity support for a defective eigenvalue

For a matrix A, we shall say the *multiplicity support* of its eigenvalue λ_* is $m \times k$ if the geometric multiplicity of λ_* is m and the Segre characteristic anchor is k. In this case,

there is a unique $X_* \in \mathbb{C}^{n\times k}$ satisfying the equation

$$\begin{aligned}(A-\lambda_* I)\,X_* &= X_*\,J_k(0)\\ C^{\mathsf{H}}\,X_* &= T\end{aligned}$$

with proper choices of $C \in \mathbb{C}^{n\times m}$ and $T \in \mathbb{C}^{m\times k}$, as we shall prove in Lemma 3.1. Here $J_k(0)$ is a nilpotent upper-triangular matrix of rank $k-1$ and can be replaced with any matrix of such kind. For integers $m, k \le n$, we define a holomorphic mapping

$$\begin{aligned}\mathbf{g} \quad : \quad \mathbb{C}^{n\times n}\times\mathbb{C}\times\mathbb{C}^{n\times k} &\longrightarrow \mathbb{C}^{n\times k}\times\mathbb{C}^{m\times k}\\ (G,\lambda,X) &\longmapsto \begin{pmatrix}(G-\lambda I)\,X - X\,S\\ C^{\mathsf{H}}\,X - T\end{pmatrix}\end{aligned} \tag{1}$$

that depends on parameters $C \in \mathbb{C}^{n\times m}$, $T \in \mathbb{C}^{m\times k}$ and an upper-triangular nilpotent matrix

$$S = \begin{bmatrix}0 & s_{12} & \cdots & s_{1k}\\ \vdots & \ddots & \ddots & \vdots\\ \vdots & & \ddots & s_{k-1,k}\\ 0 & \cdots & \cdots & 0\end{bmatrix} \quad \text{with} \quad s_{12}s_{23}\cdots s_{k-1,k} \neq 0. \tag{2}$$

We shall also denote the Jacobian and partial Jacobian

$$\begin{aligned}\mathbf{g}_{G\lambda X}(G_0,\lambda_0,X_0) &= \left.\frac{\partial \mathbf{g}(G,\lambda,X)}{\partial(G,\lambda,X)}\right|_{(G,\lambda,X)=(G_0,\lambda_0,X_0)}\\ \mathbf{g}_{\lambda X}(G_0,\lambda_0,X_0) &= \left.\frac{\partial \mathbf{g}(G,\lambda,X)}{\partial(\lambda,X)}\right|_{(G,\lambda,X)=(G_0,\lambda_0,X_0)}\end{aligned}$$

that can be considered linear transformations

$$\begin{aligned}\mathbf{g}_{G\lambda X}(G_0,\lambda_0,X_0) \quad : \quad \mathbb{C}^{n\times n}\times\mathbb{C}\times\mathbb{C}^{n\times k} &\longrightarrow \mathbb{C}^{n\times k}\times\mathbb{C}^{m\times k}\\ (G,\lambda,X) &\longmapsto \begin{pmatrix}(G-\lambda I)\,X_0 + (G_0-\lambda_0 I)\,X - X\,S\\ C^{\mathsf{H}}\,X\end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}\mathbf{g}_{\lambda X}(G_0,\lambda_0,X_0) \quad : \quad \mathbb{C}\times\mathbb{C}^{n\times k} &\longrightarrow \mathbb{C}^{n\times k}\times\mathbb{C}^{m\times k}\\ (\lambda,X) &\longmapsto \begin{pmatrix}-\lambda\,X_0 + (G_0-\lambda_0 I)\,X - X\,S\\ C^{\mathsf{H}}\,X\end{pmatrix}\end{aligned}$$

respectively. The actual matrices representing the Jacobians depend on the ordering of the bases for the domains and codomains of those linear transformations. The pseudo-inverse of a linear transformation such as $\mathbf{g}_{\lambda X}(G_0,\lambda_0,X_0)^\dagger$ is the linear transformation whose matrix representation is the pseudo-inverse of the matrix representation of $\mathbf{g}_{\lambda X}(G_0,\lambda_0,X_0)$ corresponding to the same bases.

3 Properties of the multiplicity support

The following lemma sets the foundation for the sensitivity analysis and algorithm development for a defective eigenvalue with certain multiplicity support by laying out the critical properties of the holomorphic mapping in (1).

Lemma 3.1 *Let* $A \in \mathbb{C}^{n\times n}$ *with an eigenvalue* λ_* *of multiplicity support* $m_* \times k_*$ *and* $\mathbf{g}$ *be the mapping in* (1). *The following assertions hold.*

(i) *Let* S *be as in* (2) *and*

$$T \;=\; \begin{bmatrix} 1 & \mathbf{0}_{k-1}^\top \\ \mathbf{0}_{m-1} & O \end{bmatrix}. \tag{3}$$

For almost all $C \in \mathbb{C}^{n\times m}$, *there is an* $X_* \in \mathbb{C}^{n\times k}$ *such that* $\mathbf{g}(A,\lambda_*,X_*) = \mathbf{0}$ *if and only if* $m \le m_*$ *and* $k \le k_*$. *Such an* X_* *is unique if and only if* $m = m_*$.

(ii) *Let* $m \le m_*$, $k \le k_*$ *and* S *be as in* (2). *For any* $T \in \mathbb{C}^{m\times k}$ *with* $T_{1:m,1} \neq \mathbf{0}$ *and almost all* $C \in \mathbb{C}^{n\times m}$, *assume* $\mathbf{g}(A,\lambda_*,X_*) = \mathbf{0}$. *Then* $\mathbf{g}_{G\lambda X}(A,\lambda_*,X_*)$ *is surjective, and* $\mathbf{g}_{\lambda X}(A,\lambda_*,X_*)$ *is injective if and only if* $m = m_*$ *and* $k = k_*$.

(iii) *Let* $m = m_*$, $k \le k_*$, T *be in* (3) *and columns of* $C \in \mathbb{C}^{n\times m}$ *form an orthonormal basis for the kernel of* $A - \lambda_* I$. *Then the matrix parameter* S *can be chosen so that the unique* X_* *satisfying* $\mathbf{g}(A,\lambda_*,X_*) = \mathbf{0}$ *has orthonormal columns.*

Proof. Let $N \in \mathbb{C}^{n\times m_*}$ be a matrix whose columns span the kernel $\mathcal{K}er(A-\lambda_* I)$. Then for almost all $C \in \mathbb{C}^{n\times m}$, the matrix $C^{\mathsf{H}}N$ is of full row rank if $m \le m_*$ so that there is a $\mathbf{u} \in \mathbb{C}^m$ such that

$$(C^{\mathsf{H}}N)\,\mathbf{u} \;=\; \begin{bmatrix} 1 \\ \mathbf{0}_{m-1} \end{bmatrix}$$

and $\mathbf{u}$ is unique if and only if $m = m_*$. For $m \le m_*$, let $\mathbf{x}_1 = N\,\mathbf{u}$ and assume vectors $\mathbf{x}_1,\ldots,\mathbf{x}_j \in \mathbb{C}^n$ are obtained such that $1 \le j < k$ and

$$\begin{aligned} (A-\lambda_* I)\big[\mathbf{x}_1,\cdots,\mathbf{x}_j\big] \;&=\; \big[\mathbf{x}_1,\cdots,\mathbf{x}_j\big] \begin{bmatrix} 0 & s_{12} & \cdots & s_{1j} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & s_{j-1,j} \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \\ C^{\mathsf{H}}\big[\mathbf{x}_1,\cdots,\mathbf{x}_j\big] \;&=\; \begin{bmatrix} 1 & \mathbf{0}_{j-1}^\top \\ \mathbf{0}_{m-1} & O \end{bmatrix}. \end{aligned}$$

Then the system

$$(A-\lambda_* I)\,\mathbf{x} \;=\; s_{1,j+1}\mathbf{x}_1 + \cdots + s_{j,j+1}\mathbf{x}_j$$

has an affine space solution $\big\{\mathbf{u} + N\,\mathbf{v} \;\big|\; \mathbf{v} \in \mathbb{C}^m\big\}$ and a unique solution

$$\mathbf{x}_{j+1} \;=\; \mathbf{u} - N\,(C^{\mathsf{H}}N)^{-1}C^{\mathsf{H}}\mathbf{u} \;\in\; \mathbb{C}^n$$

such that $C^{\mathsf{H}}\,\mathbf{x}_{j+1} = \mathbf{0}$ when $m = m_*$. Consequently, there is an $X_* = \big[\mathbf{x}_1,\cdots,\mathbf{x}_k\big] \in \mathbb{C}^{n\times k}$ such that (λ_*, X_*) is a solution to the system $\mathbf{g}(A,\lambda,X) = \mathbf{0}$ and X_* is unique if and only if $m = m_*$. The assertion (i) is proved.

For any parameter $T \in \mathbb{C}^{m\times k}$ of $\mathbf{g}$ with $T_{1:m,1} \neq \mathbf{0}$ and almost all $C \in \mathbb{C}^{n\times m}$, assume $\mathbf{g}(A,\lambda_*,X_*) = \mathbf{0}$. Write $X_* = \big[\mathbf{x}_1,\cdots,\mathbf{x}_k\big]$. Then $\mathbf{x}_1 \neq \mathbf{0}$ and

$$\mathbf{x}_j \;\in\; \mathcal{K}er\big((A-\lambda_* I)^j\big) \setminus \mathcal{K}er\big((A-\lambda_* I)^{j-1}\big)$$

for $j = 1, \ldots, k$. Thus X_* is of full column rank. Furthermore, by proper column permutations the matrix representation of the Jacobian $\mathbf{g}_{G\lambda X}(A, \lambda_*, X_*)$ is in the form of

$$\begin{bmatrix} X_*^\top & & & * & \cdots & * & * \\ & \ddots & & \vdots & \ddots & \vdots & \vdots \\ & & X_*^\top & * & \cdots & * & * \\ & & & C^{\mathsf{H}} & & & * \\ & & & & \ddots & & \vdots \\ & & & & & C^{\mathsf{H}} & * \end{bmatrix}$$

Therefore $\mathbf{g}_{G\lambda X}(A, \lambda_*, X_*)$ is surjective.

Let $(A, \lambda_*, \hat{X})$ be a zero of $\mathbf{g}$ and assume $m = m_*$ and $k = k_*$. Then, for almost all $C \in \mathbb{C}^{n\times m}$, the solution $\mathbf{u} = \hat{\mathbf{u}}$ of the equation $(C^{\mathsf{H}} N)\,\mathbf{u} = T_{1:m,1}$ is unique and the first column of $\hat{X}$ is

$$\hat{\mathbf{x}}_1 \;=\; N\,\hat{\mathbf{u}} \;\in\; \left(\bigcap_{j=1}^{k_*-1} \mathcal{R}an\big((A-\lambda_* I)^j\big)\right) \;\setminus\; \mathcal{R}an\big((A-\lambda_* I)^{k_*}\big) \tag{4}$$

since the smallest elementary Jordan block is $J_{k_*}(\lambda_*)$ of size $k_* \times k_*$. Assume, for certain $(\sigma, Y) \in \mathbb{C}\times\mathbb{C}^{n\times k}$, its image of the linear transformation $\mathbf{g}_{\lambda X}(A, \lambda_*, \hat{X})\big(\sigma, Y\big) = \mathbf{0}$. Namely

$$-\sigma\,\hat{X} + (A-\lambda_* I)\,Y - Y\,S \;=\; O \tag{5}$$
$$C^{\mathsf{H}}\,Y \;=\; O. \tag{6}$$

Right-multiplying both sides of the equation (5) by S yields

$$\begin{aligned} Y\,S^2 + \sigma\,\hat{X}\,S &= (A-\lambda_* I)\,Y\,S \\ &= (A-\lambda_* I)^2\,Y - \sigma\,(A-\lambda_* I)\,\hat{X} \\ &= (A-\lambda_* I)^2\,Y - \sigma\,\hat{X}\,S \end{aligned}$$

Continuing the process of recursive right-multiplying the above equation by S leads to

$$\begin{aligned} (A-\lambda_* I)^k\,Y &= Y\,S^k + k\,\sigma\,\hat{X}\,S^{k-1} \\ &= k\,\sigma\,s_{12}\,s_{23}\,\cdots\,s_{k-1,k}\big[\,O_{n\times(k-1)},\ \hat{\mathbf{x}}_1\,\big] \end{aligned}$$

with $s_{12}\,s_{23}\,\cdots\,s_{k-1,k} \neq 0$. Hence σ must be zero due to (4) and $k = k_*$. Denote columns of Y as $\mathbf{y}_1, \ldots, \mathbf{y}_k \in \mathbb{C}^n$. Then the first column of the equations (5) and (6) are

$$(A-\lambda_* I)\,\mathbf{y}_1 \;=\; \mathbf{0}, \quad C^{\mathsf{H}}\,\mathbf{y}_1 = 0$$

that implies $\mathbf{y}_1 = \mathbf{0}$. For $1 \le j < k$, using $\sigma = 0$ and $\mathbf{y}_1 = \cdots = \mathbf{y}_j = \mathbf{0}$ on the $(j+1)$-th column of the equations (5) and (6) we have $\mathbf{y}_{j+1} = \mathbf{0}$. Thus $Y = O$. As a result, $(A, \lambda_*, \hat{X})$ is a zero of $\mathbf{g}$ with injective partial Jacobian $\mathbf{g}_{\lambda X}(A, \lambda_*, \hat{X})$.

If $m < m_*$, the solution (λ_*, X_*) of $\mathbf{g}(A, \lambda, X) = \mathbf{0}$ is not isolated and thus $\mathbf{g}_{\lambda X}(A, \lambda_*, X_*)$ is not injective. Let $m = m_*$, we now prove the partial Jacobian $\mathbf{g}_{\lambda X}(A, \lambda_*, \hat{X})$ is injective

only if $k = k_*$. Assume $k < k_*$ and write $\hat{X} = \big[\, \hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_k \,\big]$. Since $\mathbf{g}(A, \lambda_*, \hat{X}) = \mathbf{0}$ and S is upper-triangular nilpotent, hence $\mathbf{x}_1 \in \mathcal{K}er(A - \lambda_* I)$. Then $k < k_*$ implies

$$\hat{\mathbf{x}}_1 \;\in\; \bigcap_{j=1}^{k+1} \mathcal{R}an\big((A - \lambda_* I)^j\big) \tag{7}$$

and there exists an $\hat{\mathbf{x}}_{k+1}$ such that $(A - \lambda_* I)\mathbf{x}_{k+1} = \hat{\mathbf{x}}_k$. For almost all $C \in \mathbb{C}^{n \times m}$, the matrix $\begin{bmatrix} A - \lambda_* I \\ C^{\mathsf{H}} \end{bmatrix}$ is of full column rank and there is a unique solution $\mathbf{z} = \mathbf{y}_1 \in \mathit{span}\{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2\}$ to the linear system

$$\begin{bmatrix} A - \lambda_* I \\ C^{\mathsf{H}} \end{bmatrix} \mathbf{z} \;=\; \begin{bmatrix} \hat{\mathbf{x}}_1 \\ \mathbf{0} \end{bmatrix}.$$

Assume we have $\mathbf{y}_1, \ldots, \mathbf{y}_j \in \mathit{span}\{\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_{j+1}\}$ for any $j \in \{1, \ldots, k-1\}$ such that

$$\begin{aligned} -\big[\, \hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_j \,\big] + (A - \lambda_* I)\big[\, \mathbf{y}_1, \cdots, \mathbf{y}_j \,\big] - \big[\, \mathbf{y}_1, \cdots, \mathbf{y}_j \,\big] S &= O \\ C^{\mathsf{H}} \big[\, \mathbf{y}_1, \cdots, \mathbf{y}_j \,\big] &= O. \end{aligned}$$

By (7), there is a unique solution $\mathbf{z} = \mathbf{y}_{j+1} \in \mathit{span}\{\mathbf{x}_1, \ldots, \hat{\mathbf{x}}_{j+1}\}$ to the linear system

$$\begin{bmatrix} A - \lambda_* I \\ C^{\mathsf{H}} \end{bmatrix} \mathbf{z} \;=\; \begin{bmatrix} \hat{\mathbf{x}}_{j+1} + s_{1,j+1}\mathbf{y}_1 + \cdots + s_{j,j+1}\mathbf{y}_j \\ \mathbf{0} \end{bmatrix}.$$

Write $Y = \big[\, \mathbf{y}_1, \cdots, \mathbf{y}_k \,\big]$. We have $\mathbf{g}_{\lambda X}(A, \lambda_*, \hat{X})\big(1, Y\big) = \mathbf{0}$ and thus the partial Jacobian $\mathbf{g}_{\lambda X}(A, \lambda_*, \hat{X})$ is not injective. As a result, the assertion (ii) is proved.

We now prove (iii). Let columns of $C \in \mathbb{C}^{n \times m}$ form an orthonormal basis for $\mathcal{K}er(A - \lambda_* I)$ and X_* be the unique matrix in assertion (i). Then C and X_* have the identical first columns. Let $X_* = Q\,R$ be a thin QR decomposition where $R = [r_{ij}]$ with $r_{11} = 1$. Since T is as in (3), it is easy to see that

$$R \;=\; \begin{bmatrix} 1 & \mathbf{0}_{k-1}^{\top} \\ \mathbf{0}_{k-1} & \hat{R} \end{bmatrix} \;\in\; \mathbb{C}^{k \times k}.$$

Thus

$$\begin{aligned} (A - \lambda_*\, I)\, Q &= Q(R\, S\, R^{-1}) \\ C^{\mathsf{H}}\, Q &= T\, R^{-1}. \end{aligned}$$

It is straightforward to verify that $T\, R^{-1} = T$ and $R\, S\, R^{-1}$ is an upper-triangular nilpotent matrix of rank $k - 1$. Consequently, the assertion (iii) is proved by setting $\check{X}$ as Q and replacing S with $R\, S\, R^{-1}$. □

4 Sensitivity of a defective eigenvalue

Lemma 3.1 reveals the critical importance of the geometric multiplicity combined with the Segre characteristic anchor of an eigenvalue and its behavior under data perturbations. In fact, the matrix and the defective eigenvalue are holomorphic functions of certain entries of the matrix as asserted in the following corollary.

Corollary 4.1 *Assume* $A \in \mathbb{C}^{n\times n}$ *and* $\lambda_* \in \mathit{eig}(A)$ *with a multiplicity support* $m \times k$. *Let the mapping* $\mathbf{g}$ *be defined in* (1) *where* $S \in \mathbb{C}^{k\times k}$ *is as in* (2), $C \in \mathbb{C}^{n\times m}$ *and* $T \in \mathbb{C}^{m\times k}$ *with* $T_{1:m,1} \neq \mathbf{0}$ *so that* $\mathbf{g}(A, \lambda_*, X_*) = \mathbf{0}$ *with an injective* $\mathbf{g}_{\lambda X}(A, \lambda_*, X_*)$. *There exist* $n^2 - m\,k + 1$ *entries of* G *forming a variable* $\mathbf{z}$ *in a neighborhood* $\Omega \ni \mathbf{z}_*$ *in* $\mathbb{C}^{n^2-m\,k+1}$, *a neighborhood* Σ *of* (A, λ_*, X_*) *in* $\mathbb{C}^{n\times n} \times \mathbb{C} \times \mathbb{C}^{n\times k}$ *and holomorphic mappings* $G : \Omega \longrightarrow \mathbb{C}^{n\times n}$, $\lambda : \Omega \longrightarrow \mathbb{C}$ *and* $X : \Omega \longrightarrow \mathbb{C}^{n\times k}$ *with* $\big(G(\mathbf{z}_*), \lambda(\mathbf{z}_*), X(\mathbf{z}_*)\big) = (A, \lambda_*, X_*)$ *such that* $\mathbf{g}(G, \lambda, X) = \mathbf{0}$ *for* $(G, \lambda, X) \in \Sigma$ *if and only if* $(G, \lambda, X) = (G(\mathbf{z}), \lambda(\mathbf{z}), X(\mathbf{z}))$ *for* $\mathbf{z} \in \Omega$.

PROOF. By Lemma 3.1, part (ii), the Jacobian $\mathbf{g}_{G\lambda X}(A, \lambda_*, X_*)$ is surjective to $\mathbb{C}^{n\times k} \times \mathbb{C}^{m\times k}$ and the partial Jacobian $\mathbf{g}_{\lambda X}(A, \lambda_*, X_*)$ is injective from $\mathbb{C} \times \mathbb{C}^{n\times k}$. Consequently, there are $m\,k-1$ entries of the variable $G \in \mathbb{C}^{n\times n}$ forming a variable $\mathbf{y}$ such that the partial Jacobian $\mathbf{g}_{\mathbf{y}\lambda X}(A, \lambda_*, X_*)$ is invertible. By the Implicit Function Theorem, the remaining entries of G excluding those forming $\mathbf{y}$ form a variable vector $\mathbf{z} \in \mathbb{C}^{n^2-m\,k+1}$ so that the assertion holds. □

Denote the collection of $n \times n$ complex matrices having an eigenvalue that shares the same multiplicity support $m \times k$ as

$$\mathcal{E}^n_{m\times k} \;:=\; \big\{\, A \in \mathbb{C}^{n\times n} \;\big|\; \text{A multiplicity support of } A \text{ is } m \times k \,\big\}. \tag{8}$$

Corollary 4.1 implies that every matrix A in $\mathcal{E}^n_{m\times k}$ has an eigenvalue λ_* along with a unique X_* such that (A, λ_*, X_*) belongs to an algebraic variety defined by the solution set of the polynomial system $\mathbf{g}(G, \lambda, X) = \mathbf{0}$.

We can now establish one of the main theorems of this paper.

Theorem 4.2 (Defective Eigenvalue Sensitivity Theorem) *The sensitivity of a defective eigenvalue is finitely bounded as long as* its *multiplicity support is preserved. More precisely, let* $A \in \mathbb{C}^{n\times n}$ *and* $\lambda_* \in \mathit{eig}(A)$ *with a multiplicity support* $m \times k$. *There is a neighborhood* Σ *of* (λ_*, A) *in* $\mathbb{C} \times \mathbb{C}^{n\times n}$ *and a neighborhood* $\Omega \in \mathbb{C}^{n^2-m\,k+1}$ *of* $\mathbf{z}_*$ *formed by* $n^2 - m\,k + 1$ *entries* A *along with holomorphic mappings* $\lambda : \Omega \to \mathbb{C}$ *and* $G : \Omega \to \mathbb{C}^{n\times n}$ *with* $(\lambda_*, A) = \big(\lambda(\mathbf{z}_*), G(\mathbf{z}_*)\big)$ *such that every* $(\tilde\lambda, \tilde A) \in \Sigma$ *with* $\tilde\lambda \in \mathit{eig}\big(\tilde A\big)$ *of multiplicity support* $m \times k$ *is equal to* $(\lambda(\mathbf{z}), G(\mathbf{z}))$ *for certain* $\mathbf{z} \in \Omega$. *Furthermore,*

$$\begin{aligned} \limsup_{\mathbf{z}\to\mathbf{z}_*} \frac{\big|\lambda(\mathbf{z}) - \lambda_*\big|}{\|G(\mathbf{z}) - A\|_F} \quad &\leq \quad \big\|\mathbf{g}_{\lambda X}(A, \lambda_*, X_*)^\dagger\big\|_2 \\ &< \quad \infty \end{aligned} \tag{9}$$

where $X_* \in \mathbb{C}^{n\times k}$ *satisfies* $\mathbf{g}(A, \lambda_*, X_*) = \mathbf{0}$ *for the mapping* $\mathbf{g}$ *defined in* (1) *that renders columns of* X_* *orthonormal.*

PROOF. The existence of neighborhoods Σ and Ω along with holomorphic mappings λ and G is a direct consequence of Corollary 4.1, which also implies

$$\mathbf{g}\big(G(\mathbf{z}), \lambda(\mathbf{z}), X(\mathbf{z})\big) \;\equiv\; O \quad \text{for all} \quad \mathbf{z} \in \Omega.$$

As a result,

$$\begin{aligned}
&\left(\left. \frac{\partial \mathbf{g}\big(G(\mathbf{z}), \lambda(\mathbf{z}), X(\mathbf{z})\big)}{\partial \mathbf{z}} \right|_{\mathbf{z}=\mathbf{z}_*} \right) (\hat{\mathbf{z}} - \mathbf{z}_*) \\
&\quad = \quad \mathbf{g}_{_G}(A, \lambda_*, X_*)\, G_{\mathbf{z}}(\mathbf{z}_*)\, (\hat{\mathbf{z}} - \mathbf{z}_*) + \mathbf{g}_{\lambda X}(A, \lambda_*, X_*) \left(\left. \frac{\partial (\lambda(\mathbf{z}), X(\mathbf{z}))}{\partial \mathbf{z}} \right|_{\mathbf{z}=\mathbf{z}_*} \right) (\hat{\mathbf{z}} - \mathbf{z}_*) \\
&\quad = \quad \mathbf{0},
\end{aligned}$$

implying

$$\begin{aligned}
\big|\lambda(\hat{\mathbf{z}}) - \lambda_*\big| \quad &\leq \quad \big\| \big(\lambda(\hat{\mathbf{z}}), X(\hat{\mathbf{z}})\big) - \big(\lambda_*, X_*\big) \big\|_2 \\
&= \quad \left\| \left. \frac{\partial (\lambda(\mathbf{z}), X(\mathbf{z}))}{\partial \mathbf{z}} \right|_{\mathbf{z}=\mathbf{z}_*} (\hat{\mathbf{z}} - \mathbf{z}_*) \right\|_2 + O\left(\big\| \hat{\mathbf{z}} - \mathbf{z}_* \big\|_2^2 \right) \\
&= \quad \big\| \mathbf{g}_{\lambda X}(A, \lambda_*, X_*)^\dagger\, \mathbf{g}_{_G}(A, \lambda_*, X_*)\, G_{\mathbf{z}}(\mathbf{z}_*)\, (\hat{\mathbf{z}} - \mathbf{z}_*) \big\|_2 + O\left(\big\| \hat{\mathbf{z}} - \mathbf{z}_* \big\|_2^2 \right) \\
&\leq \quad \big\| \mathbf{g}_{\lambda X}(A, \lambda_*, X_*)^\dagger \big\|_2 \, \big\| G(\mathbf{z}) - A_* \big\|_F + O\left(\big\| \hat{\mathbf{z}} - \mathbf{z}_* \big\|_2^2 \right)
\end{aligned}$$

since, by proper row and column permutations, the partial Jacobian $\mathbf{g}_{_G}(A, \lambda_*, X_*)$ has a matrix representation

$$\begin{bmatrix} X_*^\top & & \\ & \ddots & \\ & & X_*^\top \end{bmatrix}$$

with a unit 2-norm due to orthonormal columns of X_*, leading to the inequality (9). The norm $\big\| \mathbf{g}_{\lambda X}(A, \lambda_*, X_*)^\dagger \big\|_2$ is finite because $\mathbf{g}_{\lambda X}(A, \lambda_*, X_*)$ is injective by Lemma 3.1. □

In light of Theorem 4.2, we can introduce an $m \times k$ *condition number* of an eigenvalue $\lambda_* \in \mathit{eig}(A)$ as

$$\begin{aligned}
\tau_{A,m\times k}(\lambda_*) \quad &:= \quad \inf_{C,T,S} \left\| \big(\mathbf{g}_{\lambda X}(A, \lambda_*, X_*)\big)^\dagger \right\| \\
&< \quad \infty
\end{aligned} \tag{10}$$

where $\mathbf{g}$ is the mapping defined in (1) and the infimum is taken over all the proper choices of matrix parameters C, T, S that render the columns of the unique X_* orthonormal so that $\mathbf{g}(A, \lambda_*, X_*) = \mathbf{0}$. We shall refer to $\tau_{A,m\times k}(\lambda_*)$ as the *multiplicity support condition number* if the specific m and k are irrelevant in the discussion. From Lemma 3.1, the $m \times k$ condition number is infinity *only if* either m is less than the actual geometric multiplicity or k is less than the Segre characteristic anchor. Consequently, the multiplicity support condition number $\tau_{A,m\times k}(\lambda_*)$ is large only if A is close to a matrix $\tilde{A}$ that possesses an eigenvalue $\tilde{\lambda} \approx \lambda_*$ whose geometric multiplicity is larger than m or its Segre characteristic anchor is larger than k.

We can now revisit the question:

Is a defective eigenvalue hypersensitive to perturbations?

The answer is not as simple as the question may seem to be.

It is well documented in the literature that, under an *arbitrary* perturbation ΔA on the matrix A, a defective eigenvalue of A generically disperses into a cluster of simple eigenvalues with an error bound proportional to $\|\Delta A\|_2^{\frac{1}{l}}$ where l is the size of the largest Jordan block associated with the eigenvalue [2, p. 58][3, 13]. Similar and related sensitivity results can be found in the works such as [1, 14, 15]. This error bound implies that the asymptotic sensitivity of a defective eigenvalue is infinity, and only a fraction $\frac{1}{l}$ of the data accuracy passes on to the accuracy of the eigenvalue. For instance, if the largest Jordan block is 5×5, only three correct digits can be expected from the computed eigenvalues regarding the defective eigenvalue since one fifth the hardware precision (about 16 digits) remains in the forward accuracy. It is also known that the mean of the cluster emanating from the defective eigenvalue under perturbations is not hypersensitive [10, 17].

Kahan is the first to discover the finite sensitivity $\frac{1}{m}\|P\|_2$ of a multiple eigenvalue under constrained perturbations that preserve the algebraic multiplicity m, where P is the spectral projector associated with the eigenvalue. This spectral projector norm is large only if a small perturbation on the matrix can increase the multiplicity [9]. As pointed out by Kahan, the seemingly infinite sensitivity of a defective eigenvalue may not be a conceptually meaningful measurement for the condition of a *multiple eigenvalue* since arbitrary perturbations do not maintain the characteristics of the eigenvalue as being multiple. Theorem 4.2 sheds light on another intriguing and pleasant property of a defective eigenvalue: Its algebraic multiplicity does *not* need to be maintained under data perturbations for its sensitivity to be under control, as long as the geometric multiplicity *and* the Segre characteristic anchor are preserved. As a result, the condition number $\tau_{A,m\times k}(\lambda_*)$ provides a new and different measurement on the sensitivity of a *defective* eigenvalue λ_* when its multiplicity support is preserved. More importantly, a defective eigenvalue can be accurately computed in numerical computation from imposing the constraints on the multiplicity support, as we shall demonstrate in later sections.

Interestingly, the same eigenvalue can be ill-conditioned in the spectral projector norm while being well conditioned in multiplicity support and vice versa (c.f. Example 4 in §11), no contradiction whatsoever.

Even if perturbations are unconstrained, the problem of computing a defective eigenvalue may not have to be hypersensitive at all if the problem is properly generalized, i.e. regularized. We shall prove in Theorem 6.1 that the $m \times k$ condition number still provides the finitely bounded sensitivity of λ_* as what we call the $m \times k$ pseudo-eigenvalue of A, and this condition number is large only if m or k can be increased by small perturbations.

There are further subtleties on the condition of a defective eigenvalue. The sensitivity is finitely bounded if the multiplicity or the multiplicity support of *the* eigenvalue is preserved. The set $\mathcal{E}^n_{m\times k}$ in (8) is not a manifold so the Tubular Neighborhood Theorem does not apply. As a result, requiring the matrix to maintain a multiplicity support $m \times k$ is not enough to dampen the sensitivity of a particular defective eigenvalue with that multiplicity support. The matrix staying on $\mathcal{E}^n_{m\times k}$ does not guarantee the finite sensitivity of a defective

eigenvalue. If a matrix $A \in \mathcal{E}^n_{m\times k}$ has two eigenvalues of the same multiplicity support $m \times k$, then A is in the intersection of images of two holomorphic mappings described in Corollary 4.1. When A drifts on $\mathcal{E}^n_{m\times k}$, the multiplicity support $m \times k$ may be maintained for one eigenvalue but lost on the other. Consequently, the "other" defective eigenvalue still disperses into a cluster.

5 Backward errors near a defective eigenvalue

Lemma 3.1 makes it possible to calculate a defective eigenvalue in numerical computation. When an approximate value $\tilde\lambda$ of the defective eigenvalue is obtained as a component of the approximate zero $(A, \tilde\lambda, \tilde X)$ of $\mathbf{g}$ in (1), a backward error measure of $\tilde\lambda$ can be easily obtained from the following lemma.

Corollary 5.1 *Assume $A \in \mathbb{C}^{n\times n}$ and the mapping $\mathbf{g}$ is defined in (1) with any parameters $C \in \mathbb{C}^{n\times m}$, S as in (2) and $T \in \mathbb{C}^{m\times k}$ with $T_{1:m,1} \neq \mathbf{0}$. For any $\big(\tilde\lambda, \tilde X\big) \in \mathbb{C}\times\mathbb{C}^{n\times k}$ with a full-ranked $\tilde X$ and residual $(A-\tilde\lambda I)\,\tilde X - \tilde X\,S = E$, the backward error bound of $\tilde\lambda$ is $\|E\|_2 \left\|\tilde X^\dagger\right\|_2$ in the sense that $\tilde\lambda$ is an exact eigenvalue of $A - E\,\tilde X^\dagger$ associated with an elementary Jordan block $J_l(\tilde\lambda)$ of size $l \geq k$, and the bound is $\|E\|_2$ if columns of $\tilde X$ are orthonormal.*

Proof. Since $\tilde X$ is of full column rank, we have $\tilde X^\dagger\,\tilde X = I$ and thus $E = E\,\tilde X^\dagger\,\tilde X$, leading to $\big(A - E\,\tilde X^\dagger - \tilde\lambda I\big)\,\tilde X = \tilde X\,S$ and the assertions follow from a straightforward verification. □

Corollary 5.1 provides a backward error on the approximate eigenvalue $\tilde\lambda$ in the sense that it is an exact defective eigenvalue of the matrix $A - E\,\tilde X^\dagger$. The backward error bound $\|E\|_2 \left\|\tilde X^\dagger\right\|_2$ points to the importance of orthonormalizing columns of $\tilde X$ in achieving highest possible accuracy. Lemma 3.1 ensures that we can choose the parameter C and S in (1) so that columns of $\tilde X$ are orthonormal or nearly orthonormal. In that way, the magnitude $\|E\|_2$ of the the residual $E = (A - \tilde A I)\,\tilde X - \tilde X\,S$ directly measures such a backward error. The method of orthonormalization is elaborated in §10 along with numerical results confirming the accuracy improvement.

There is another way to measure the backward error so that the computed eigenvalue is an exact eigenvalue of a nearby matrix having an identical multiplicity support, as asserted in the following lemma.

Corollary 5.2 *Let $A \in \mathbb{C}^{n\times n}$ and $\lambda_* \in \mathit{eig}(A)$ with a multiplicity support $m \times k$. As parameters of $\mathbf{g}$ in (1), assume $S \in \mathbb{C}^{k\times k}$ is as in (2), $T \in \mathbb{C}^{m\times k}$ is as in (3) and $C \in \mathbb{C}^{n\times m}$ is full-ranked such that the mapping $\mathbf{g}(A, \lambda_*, X)$ has a unique zero X_* and $\mathbf{g}_{\lambda X}(A, \lambda_*, X_*)$ is injective. Further assume $\tilde A \in \mathbb{C}^{n\times n}$ represents the available data of A with sufficiently small error $\|A - \tilde A\|_2$. Then every sufficiently accurate approximation $\tilde\lambda$ of λ_* is an exact eigenvalue of $\tilde A + \Delta\tilde A$ with the identical multiplicity support $m \times k$ and*

$$\left\|\Delta\tilde A\right\|_2 \;=\; \left\|\left[E_1 - (\tilde A - \tilde\lambda I)\,C^{\mathsf{H}\dagger}\,E_2 + C^{\mathsf{H}\dagger}\,E_2\,S\right](\tilde X - C^{\mathsf{H}\dagger}\,E_2)^\dagger\right\|_2 \tag{11}$$

where $\tilde{X}$ *is the least squares solution of* $\mathbf{g}(A, \tilde{\lambda}, X) = \mathbf{0}$ *with the residual*

$$\mathbf{g}(\tilde{A}, \tilde{\lambda}, \tilde{X}) \;=\; \begin{bmatrix} (\tilde{A} - \tilde{\lambda} I)\,\tilde{X} - \tilde{X}\,S \\ C^{\mathsf{H}}\,\tilde{X} - T \end{bmatrix} \;=\; \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}$$

PROOF. Since $C \in \mathbb{C}^{n\times m}$ is of full rank and $C^{\mathsf{H}}\,\tilde{X} - T \,=\, E_2 \,=\, C^{\mathsf{H}}\,C^{\mathsf{H}\dagger}\,E_2$, we have

$$C^{\mathsf{H}}\left(\tilde{X} - C^{\mathsf{H}\dagger}\,E_2\right) \;=\; T$$

where $C^{\mathsf{H}\dagger}$ is the pseudo-inverse of C^{H}. Since X_* and $\tilde{X}$ are least squares solutions of linear systems $\mathbf{g}(A, \lambda_*, X) = \mathbf{0}$ and $\mathbf{g}(\tilde{A}, \tilde{\lambda}, X) = \mathbf{0}$ respectively, the norms $\|X_* - \tilde{X}\|_2$, $\|E_1\|_2$ and $\|E_2\|_2$ are sufficiently small so that $\tilde{X} - C^{\mathsf{H}\dagger}\,E_2$ is of full column rank. Let

$$\begin{aligned} (\tilde{A} - \tilde{\lambda} I)\big(\tilde{X} - C^{\mathsf{H}\dagger}\,E_2\big) - \big(\tilde{X} - C^{\mathsf{H}\dagger}\,E_2\big)\,S \;&=\; E_3 \\ &\equiv\; E_3\,\big(\tilde{X} - C^{\mathsf{H}\dagger}\,E_2\big)^{\dagger}\big(\tilde{X} - C^{\mathsf{H}\dagger}\,E_2\big). \end{aligned}$$

Namely

$$\big[\big(\tilde{A} + \Delta\tilde{A}\big) - \tilde{\lambda} I\big]\,\hat{X} \;=\; \hat{X}\,S$$

with

$$\hat{X} \;=\; \tilde{X} - C^{\mathsf{H}\dagger}\,E_2 \quad \text{and} \quad \Delta\tilde{A} \;=\; -\,E_3\,\hat{X}^{\dagger}$$

that satisfies (11), leading to $\mathbf{g}(\tilde{A} + \Delta\tilde{A}, \tilde{\lambda}, \hat{X}) = \mathbf{0}$.

For sufficiently small $|\lambda_* - \tilde{\lambda}|$ and $\|A - \tilde{A}\|_2$, the norms $\|E_1\|_2$, $\|E_2\|_2$, $\|E_3\|_2$ and $\|\Delta\tilde{A}\|_2$ are all small. Since $T_{2:m,1} = \mathbf{0}$, hence

$$\begin{bmatrix} (\tilde{A} + \Delta\tilde{A}) - \tilde{\lambda} I \\ (C_{1:n,2:m})^{\mathsf{H}} \end{bmatrix} \hat{X}_{1:n,1} \;=\; \mathbf{0}$$

with $\hat{X}_{1:n,1} \neq \mathbf{0}$, implying both $\begin{bmatrix} (\tilde{A} + \Delta\tilde{A}) - \tilde{\lambda} I \\ (C_{1:n,2:m})^{\mathsf{H}} \end{bmatrix}$ and $\begin{bmatrix} A + \lambda I \\ (C_{1:n,2:m})^{\mathsf{H}} \end{bmatrix}$ are rank deficient with $\big\|(\tilde{A} + \Delta\tilde{A} - \tilde{\lambda} I) - (A - \lambda I)\big\|$ sufficiently small. Thus

$$\textit{nullity}\left(\,(\tilde{A} + \Delta\tilde{A}) - \tilde{\lambda}\,I\,\right) \;=\; \textit{nullity}\,(\,A - \lambda_* I\,) \;=\; m$$

Namely the geometric multiplicity of $\tilde{\lambda}$ as an exact eigenvalue of $\tilde{A} + \Delta\tilde{A}$ is m.

Since $\mathbf{g}(A, \lambda_*, X_*) = \mathbf{0}$ and $\mathbf{g}_{\lambda X}(A, \lambda_*, X_*)$ is injective, the first column of X_* is not in the kernel $\mathcal{K}\!er\big((A - \lambda_*\,I)^{k+1}\big)$. Consequently, the first column $\hat{X}_{1:n,1}$ of $\hat{X}$ does not belong to the kernel $\mathcal{K}\!er\big((\tilde{A} + \Delta\tilde{A} - \tilde{\lambda}\,I)^{k+1}\big)$ and $\mathbf{g}_{\lambda X}(\tilde{A} + \Delta\tilde{A}, \tilde{\lambda}, \hat{X})$ is injective when $|\lambda_* - \tilde{\lambda}|$ and $\|A - \tilde{A}\|_2$ are sufficiently small, implying the Segre characteristic anchor of $\tilde{\lambda} \in \textit{eig}(\tilde{A} + \Delta\tilde{A})$ is exactly k. □

When a matrix A possesses an eigenvalue λ_* with a multiplicity support $m \times k$ but is known through empirical data in $\tilde{A}$, solving the system $\mathbf{g}(\tilde{A}, \lambda, X) = \mathbf{0}$ for its nonlinear least squares solution $(\tilde{\lambda}, \tilde{A})$ means finding an exact eigenvalue $\tilde{\lambda}$ of a nearby matrix

$\tilde{A} + \Delta\tilde{A}$ whose multiplicity support is identical to that of the underlying hidden matrix A and the backward error bound on $\Delta\tilde{A}$ is proportional to the residual $\left\|\mathbf{g}(\tilde{A}, \tilde{\lambda}, \tilde{X})\right\|_F$ as shown in (11). Minimizing the residual reduces, if not minimizes, the distance from $\tilde{A}$ to the set $\mathcal{E}^n_{m\times k}$ of matrices with a multiplicity support $m \times k$. As a result, Theorem 4.2 ensures the computed eigenvalue $\tilde{\lambda}$ is accurate in the sense that the forward error $|\tilde{\lambda} - \lambda_*|$ is in the order of the data error $\left\|A - \tilde{A}\right\|_2$.

Requiring $|\tilde{\lambda} - \lambda_*|$ to be sufficiently small is to ensure $\tilde{\lambda}$ approximates the correct eigenvalue λ_* in case there are two or more eigenvalues of A sharing the same multiplicity support.

6 A well-posed defective eigenvalue problem

The problem of finding an eigenvalue of a matrix in its conventional meaning is ill-posed when the eigenvalue is multiple and defective because the sensitivity of the eigenvalue is infinite with respect to arbitrary perturbations on the matrix. Lacking a Lipschitz continuity of an eigenvalue with respect to data, such a problem is not suitable for numerical computation unless the problem is properly modified, or better known as being *regularized*.

We can alter the problem of

finding an eigenvalue of a matrix A

to

finding a $\tilde{\lambda}$ *so that* $(\tilde{\lambda}, \tilde{X})$ *is local least squares solution to* $\mathbf{g}(A, \lambda, X) = \mathbf{0}$

where $\mathbf{g}$ is the mapping defined in (1) with proper parameters. When the matrix A has an eigenvalue λ_* of multiplicity support $m \times k$, there is an X_* such that (λ_*, X_*) is an exact solution to $\mathbf{g}(A, \lambda, X) = \mathbf{0}$ with zero residual. However, when A is known through its empirical data in $\tilde{A}$, a local least squares solution $(\tilde{\lambda}, \tilde{X})$ to the system $\mathbf{g}(\tilde{A}, \lambda, X) = \mathbf{0}$ generally has a residual $\left\|\mathbf{g}(\tilde{A}, \tilde{\lambda}, \tilde{X})\right\|_2 > 0$, and $\tilde{\lambda}$ is not an eigenvalue of either A or $\tilde{A}$. For the convenience of elaboration, we call such a $\tilde{\lambda}$ an $m \times k$ *pseudo-eigenvalue* of $\tilde{A}$.

Theorem 6.1 (Pseudo-Eigenvalue Theorem) *Let* λ_* *be an eigenvalue of* $A \in \mathbb{C}^{n\times n}$ *with a multiplicity support* $m \times k$ *along with* $X_* \in \mathbb{C}^{n\times k}$ *satisfying* $\mathbf{g}(A, \lambda_*, X_*) = \mathbf{0}$ *with a properly defined mapping* $\mathbf{g}$ *as in* (1). *The following assertions hold.*

(i) *The exact eigenvalue* λ_* *of* A *is an* $m \times k$ *pseudo-eigenvalue of* A.

(ii) *There are neighborhoods* Φ *of* A *in* $\mathbb{C}^{n\times n}$ *and* Λ *of* λ_* *in* $\mathbb{C}$ *respectively such that every matrix* $\tilde{A} \in \Phi$ *has a unique* $m \times k$ *pseudo-eigenvalue* $\tilde{\lambda} \in \Lambda$ *that is Lipschitz continuous with respect to* $\tilde{A}$.

(iii) *For every matrix* $\check{A} \in \Phi$ *serving as empirical data of* A, *there is a unique* $m \times k$ *pseudo-eigenvalue* $\check{\lambda} \in \Lambda$ *of* $\check{A}$ *such that*

$$\left|\check{\lambda} - \lambda_*\right| \quad \leq \quad \tau_{A,m\times k}(\lambda_*) \left\|\check{A} - A\right\|_2 + O\left(\left\|\check{A} - A\right\|_2^2\right) \tag{12}$$

where $\tau_{A,m\times k}(\lambda_*)$ *is the* $m \times k$ *condition number of* $\lambda_* \in \mathit{eig}(A)$ *defined in* (10).

PROOF. The assertion (i) is a result of Lemma 3.1, part (i). Let Ψ be a small neighborhood of (λ_*, X_*) in $\mathbb{C} \times \mathbb{C}^{n\times k}$ so that $\{A\} \times \overline{\Psi}$ is a subset of Σ in Corollary 4.1. To prove the existence of a pseudo-eigenvalue of any matrix near A, assume there is a matrix $\tilde{A}$ with $\|\tilde{A} - A\|_2 < \varepsilon$ for any ε such that $\min_{(\lambda,X)\in\overline{\Psi}} \left\|\mathbf{g}(\tilde{A}, \lambda, X)\right\|_2$ are not attainable in Ψ. Let $\varepsilon \to 0$. Then $\tilde{A} \to A$ and there exists an $(\hat{\lambda}, \hat{X}) \in \overline{\Psi} \setminus \Psi$ such that

$$\mathbf{g}(A, \hat{\lambda}, \hat{X}) \;=\; \min_{(\lambda,X)\in\overline{\Psi}} \left\|\mathbf{g}(A, \lambda, X)\right\|_2 \;=\; 0$$

and $\hat{\lambda} \neq \lambda(\mathbf{z}_*)$. This is a contradiction to Corollary 4.1. As a result, there are neighborhoods Φ of A and Λ of λ_* respectively such that every matrix in Φ has an $m \times k$ pseudo-eigenvalue in Λ. By the local convergence of the Gauss-Newton iteration, we can assume Φ and Ψ are sufficiently small so that the residual $\|\mathbf{g}(G, \lambda, X)\|_2$ is small for all $G \in \Phi$ and $(\lambda, X) \in \Psi$. Consequently, for every $\tilde{A} \in \Phi$, the Gauss-Newton iteration converges to a local minimum point $(\tilde{\lambda}, \tilde{X})$ of $\|\mathbf{g}(\tilde{A}, \lambda, X)\|_2$ from any initial iterate $(\lambda_0, X_0) \in \Psi$. As a result, this local minimum points is unique in Ψ and is thus the absolute minimum point in Ψ. To prove the Lipschitz continuity of the pseudo-eigenvalue, let $\tilde{A}, \check{A} \in \Phi$ with corresponding minimum points $(\tilde{\lambda}, \tilde{X})$ and $(\check{\lambda}, \check{X})$ of $\|\mathbf{g}\|_2$ in Ψ. The Gauss-Newton iteration from $(\tilde{\lambda}, \tilde{X})$ toward $(\check{\lambda}, \check{X})$ in a single step

$$(\lambda_1, X_1) \;=\; (\tilde{\lambda}, \tilde{X}) - \mathbf{g}_{\lambda X}(\check{A}, \tilde{\lambda}, \tilde{X})^\dagger \, \mathbf{g}(\check{A}, \tilde{\lambda}, \tilde{X})$$

yields $\left\|(\lambda_1, X_1) - (\check{\lambda}, \check{X})\right\|_2 \le \mu \left\|(\check{\lambda}, \check{X}) - (\tilde{\lambda}, \tilde{X})\right\|_2$ with $0 \le \mu < 1$. Using the identity $(\tilde{\lambda}, \tilde{X}) = (\tilde{\lambda}, \tilde{X}) - \mathbf{g}_{\lambda X}(\tilde{A}, \tilde{\lambda}, \tilde{X})^\dagger \, \mathbf{g}(\tilde{A}, \tilde{\lambda}, \tilde{X})$, there is a constant γ such that

$$\begin{aligned}
\left\|(\check{\lambda}, \check{X}) - (\tilde{\lambda}, \tilde{X})\right\|_2 \;&\le\; \frac{1}{1-\mu} \left\|(\lambda_1, X_1) - (\tilde{\lambda}, \tilde{X})\right\|_2 \\
&\le\; \frac{1}{1-\mu} \Big(\left\|\mathbf{g}_{\lambda X}(\tilde{A}, \tilde{\lambda}, \tilde{X})^\dagger\right\|_2 \left\|\mathbf{g}(\tilde{A}, \tilde{\lambda}, \tilde{X}) - \mathbf{g}(\check{A}, \tilde{\lambda}, \tilde{X})\right\|_2 \\
&\qquad + \left\|\mathbf{g}_{\lambda X}(\tilde{A}, \tilde{\lambda}, \tilde{X})^\dagger - \mathbf{g}_{\lambda X}(\check{A}, \tilde{\lambda}, \tilde{X})^\dagger\right\|_2 \left\|\mathbf{g}(\tilde{A}, \tilde{\lambda}, \tilde{X})\right\|_2 \Big) \\
&\le\; \gamma \left\|\tilde{A} - \check{A}\right\|_2
\end{aligned}$$

for all $\tilde{A}, \check{A} \in \Phi$. Namely, the $m \times k$ pseudo-eigenvalue is Lipschitz continuous with respect to the matrix. Furthermore, by setting $(\tilde{A}, \tilde{\lambda}, \tilde{X}) = (A, \lambda_*, X_*)$ in the above inequalities we have (12) because the residual $\left\|\mathbf{g}(A, \lambda_*, X_*)\right\|_2 = 0$ and thus $\mu = 0$. □

The above Pseudo-Eigenvalue Theorem establishes a rigorous and thorough regularization of the ill-posed problem in computing a defective eigenvalue so that the problem of computing a pseudo-eigenvalue enjoys unique existence and Lipschitz continuity of the solution that approximates the underlying defective eigenvalue with an error bound proportional to the data error. Furthermore, this theorem reaffirms the $m \times k$ condition number as a bona fide sensitivity measure of defective eigenvalues.

This regularization makes it possible to compute defective eigenvalues accurately using floating point arithmetic even if the matrix data are perturbed, and we shall present such an algorithm in next section.

7 An algorithm for computing a defective eigenvalue

The Pseudo-eigenvalue Theorem sets the foundation for accurate computation of a defective eigenvalue even if the the matrix data are empirical, provided that the multiplicity support can be obtained. The computation is under the assumptions that the given matrix A is the data representation of a underlying matrix possessing a defective eigenvalue and an initial estimate λ_0 is close to that eigenvalue. Assuming the multiplicity support $m \times k$ is known, identified or estimated, we also need to set up the matrix parameters $C \in \mathbb{C}^{n\times m}$ and $S \in \mathbb{C}^{k\times k}$, while using the parameter $T \in \mathbb{C}^{m\times k}$ as in (3). To elaborate the strategy for setting up C, we consider $\lambda_* \in \mathit{eig}(A)$ with a multiplicity support $m \times k$. The requirements for the parameter C are two-fold:

(i) The parameter matrix C of $\mathbf{g}$ in (1) must be generic enough so that the matrix $\begin{bmatrix} A - \lambda_* I \\ C^{\mathsf{H}} \end{bmatrix}$ is injective and $\begin{bmatrix} A - \lambda_0 I \\ (C_{1:n,2:m})^{\mathsf{H}} \end{bmatrix}$ is rank-deficient.

(ii) The parameter C must be generic enough so that the linear system

$$\begin{bmatrix} A - \lambda_* I \\ C^{\mathsf{H}} \end{bmatrix} \mathbf{z} \;=\; \begin{bmatrix} \mathbf{0} \\ T_{1:m,1} \end{bmatrix} \tag{13}$$

has a unique solution that does not belong to the kernel $\mathcal{K}\mathit{er}\big((\tilde{A} - \tilde{\lambda}\, I)^{k+1}\big)$.

Let $N \in \mathbb{C}^{n\times m}$ be a matrix whose columns form an orthonormal basis for the kernel $\mathcal{K}\mathit{er}(A - \lambda_* I)$. A naive choice would be setting $C = N$, which certainly satisfies the requirement (i), and the equation (13) yields the solution $\mathbf{z}$ that is identical to the first column of N and may be closed to $\mathcal{K}\mathit{er}\big((\tilde{A} - \tilde{\lambda}\, I)^{k+1}\big)$. Instead, let $\mathbf{d} = Q \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix}$ be a random unit column vector and its QR decomposition. Then $\mathbf{z} = N\,\mathbf{d}$ is a random vector in the kernel of $A - \lambda_* I$ and, for almost all such random vector $\mathbf{d}$, does not belong to the kernel $\mathcal{K}\mathit{er}\big((\tilde{A} - \tilde{\lambda}\, I)^{k+1}\big)$. Thus we can set $C = N\,Q$ so that the solution $\mathbf{z}$ of the equation (13) equals to $N\,\mathbf{d}$.

Of course, we do not have the exact value of λ_* and we likely do not have the exact data for A either. We can assume the data are reasonably accurate and the initial estimate λ_0 is close to λ_*. We can practically ensure these requirements of C to be satisfied by the following simple process: Let $N \in \mathbb{C}^{n\times m}$ be the matrix whose columns form an orthonormal basis for the numerical kernel $\mathcal{K}\mathit{er}_\theta(A - \lambda_0 I)$ of dimension m. Namely, columns of N span the vector space spanned by the last m right singular vectors of $A - \lambda_0 I$. Such an N can be computed by a rank-revealing method such as [12] or a singular value decomposition. Set $C = N\,Q$ where $Q\,R = \mathbf{d}$ is the QR decomposition of a random vector $\mathbf{d}$.

With C and T already available, we can set up

$$\mathbf{x}_1^{(0)} \quad = \quad C \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} \tag{14}$$

$$\mathbf{x}_{j+1}^{(0)} \quad = \quad \alpha_j \begin{bmatrix} A - \lambda_0 I \\ C^{\mathsf{H}} \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{x}_j^{(0)} \\ \mathbf{0} \end{bmatrix} \quad \text{for} \quad j = 1, \ldots, k-1 \tag{15}$$

$$S \quad = \quad \left[\begin{array}{c|ccc} 0 & \alpha_1 & & \\ \vdots & & \ddots & \\ 0 & & & \alpha_{k-1} \\ \hline 0 & 0 & \cdots & 0 \end{array}\right] \tag{16}$$

where, for $j = 1, \ldots, k-1$, the scalar α_j scales $\mathbf{x}_{j+1}^{(0)}$ into a unit vector. Denote $X_0 = \left[\mathbf{x}_1^{(0)}, \cdots, \mathbf{x}_k^{(0)}\right]$. Then $\mathbf{g}(A, \lambda_0, X_0) \approx \mathbf{0}$ and we can apply the Gauss-Newton iteration

$$(\lambda_{j+1}, X_{j+1}) \quad = \quad (\lambda_j, X_j) - \mathbf{g}_{\lambda X}(A, \lambda_j, X_j)^{\dagger}\, \mathbf{g}(A, \lambda_j, X_j) \quad \text{for} \quad j = 0, 1, \ldots \tag{17}$$

that converges to (λ_*, X_*) assuming the initial estimate λ_0 is sufficiently close to λ_*. When the iteration stops at the j-th step, a QR decomposition of the matrix representing $\mathbf{g}_{\lambda X}(A, \lambda_j, X_j)$ is available and thus an estimate $\left\|\mathbf{g}_{\lambda X}(A, \lambda_j, X_j)^{\dagger}\right\|_2$ of the $m \times k$ condition number can be computed by a couple of steps of inverse iteration [12] with a negligible cost.

A pseudo-code of Algorithm EigenIteration is given in Fig. 2.

8 Taking advantage of the Jacobian structure

The main computing cost of Algorithm EigenIteration occurs at solving the linear system

$$\mathbf{g}_{\lambda X}(A, \lambda_j, X_j)(\sigma, Y) \quad = \quad \mathbf{g}(A, \lambda_j, X_j)$$

where the partial Jacobian $\mathbf{g}_{\lambda X}(A, \lambda_j, X_j)$ is of dimensions $(n\,k + m\,k) \times (n\,k + 1)$ and its QR decomposition is needed. The partial Jacobian is pleasantly structured with a proper arrangement so that the cost of QR decomposition can be reduced substantially.

Let $X = \left[\mathbf{x}_1, \cdots, \mathbf{x}_k\right]$, the image $\mathbf{g}(G, \lambda, X) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{m \times k}$ of the mapping $\mathbf{g}$ can be arranged as

$$\mathbf{g}(G, \lambda, \left[\mathbf{x}_1, \cdots, \mathbf{x}_k\right]) \quad =$$

$$\begin{bmatrix}
C^{\mathsf{H}}\mathbf{x}_k & & & & & -T_{1:m,k} \\
(G - \lambda I)\,\mathbf{x}_k & -s_{k-1,k}\,\mathbf{x}_{k-1} & -s_{k-2,k}\,\mathbf{x}_{k-2} & -\cdots & -s_{1k}\,\mathbf{x}_1 & \\
 & C^{\mathsf{H}}\mathbf{x}_{k-1} & & & & -T_{1:m,k-1} \\
 & (G - \lambda I)\,\mathbf{x}_{k-1} & -s_{k-2,k-1}\,\mathbf{x}_{k-2} & -\cdots & -s_{1,k-1}\,\mathbf{x}_1 & \\
 & \ddots & & \ddots & \vdots & \vdots \\
 & & \ddots & & \vdots & \vdots \\
 & & & \ddots & -s_{12}\,\mathbf{x}_1 & \vdots \\
 & & & & C^{\mathsf{H}}\mathbf{x}_1 & -T_{1:m,1} \\
 & & & & (G - \lambda\, I)\,\mathbf{x}_1 &
\end{bmatrix}.$$

Algorithm EigenIteration

Input: matrix A, initial eigenvalue estimate λ_0, geometric multiplicity m, Segre characteristic anchor k.

- obtain the last m right singular vectors of $A-\lambda_0 I$ to form $N \in \mathbb{C}^{n\times m}$
- obtain a random vector $\mathbf{d} \in \mathbb{C}^m$ and its QR decomposition $\mathbf{d} = Q\,R$.
- set $C = N\,Q$ and $\mathbf{x}_1^{(0)}$ as its first column.
- for $j = 1, 2, \ldots, k-1$ do

  solve $\begin{bmatrix} A-\lambda_0 I \\ C^{\mathsf{H}} \end{bmatrix} \mathbf{u} = \begin{bmatrix} \mathbf{x}_j^{(0)} \\ \mathbf{0} \end{bmatrix}$ for the least squares solution $\mathbf{u}$.

  set $\mathbf{x}_{j+1}^{(0)} = \alpha_j\,\mathbf{u}$ so that $\left\|\mathbf{x}_{j+1}^{(0)}\right\|_2 = 1$.

  end do
- set $X_0 = \left[\,\mathbf{x}_1^{(0)}, \cdots, \mathbf{x}_k^{(0)}\,\right]$ and set S as in (16)
- set the mapping $\mathbf{g}$ as in (1) using C and S
- for $j = 0, 1, \ldots$ do
  - solve the linear system $\mathbf{g}_{\lambda X}(A, \lambda_j, X_j)\,(\sigma, Y) = \mathbf{g}(A, \lambda_j, X_j)$ for the least squares solution (σ, Y)
  - set $\lambda_{j+1} = \lambda_j - \sigma$, $X_{j+1} = X_j - Y$.
  - if $\left\|\mathbf{g}(A, \lambda_j, X_j)\right\|_2 < \left\|\mathbf{g}(A, \lambda_{j+1}, X_{j+1})\right\|_2$ then

    set $(\hat\lambda, \hat X) = (\lambda_j, X_j)$, break the loop. end if

  end do

Output: eigenvalue $\hat\lambda$, backward error bound $\left\|\mathbf{g}(A, \hat\lambda, \hat X)\right\|_2 \left\|X^\dagger\right\|_2$, $m \times k$ condition number $\left\|\mathbf{g}_{\lambda X}(A, \hat\lambda, \hat X)^\dagger\right\|_2$

Figure 2: Algorithm EigenIteration

As a result, the partial Jacobian matrix in a blockwise upper-triangular form

$$\frac{\partial \mathbf{g}(G, \lambda, X)}{\partial(\mathbf{x}_k, \ldots, \mathbf{x}_1, \lambda)} = \begin{bmatrix}
C^{\mathsf{H}} & O & O & \cdots & O & \mathbf{0} \\
G-\lambda I & -s_{k-1,k}\,I & -s_{k-2,k}\,I & \cdots & -s_{1k}\,I & -\mathbf{x}_k \\
 & C^{\mathsf{H}} & O & \cdots & O & \mathbf{0} \\
 & G-\lambda I & -s_{k-2,k-1}\,I & \cdots & -s_{1,k-1}\,I & -\mathbf{x}_{k-1} \\
 & & \ddots & \ddots & \vdots & \vdots \\
 & & & \ddots & \vdots & \vdots \\
 & & & & C^{\mathsf{H}} & O & \mathbf{0} \\
 & & & & G-\lambda I & -s_{12}\,I & -\mathbf{x}_2 \\
 & & & & & C^{\mathsf{H}} & \mathbf{0} \\
 & & & & & G-\lambda I & -\mathbf{x}_1
\end{bmatrix}.$$

We can further assume the matrix G is already reduced to a Hessenberg form. Then the matrix block

$$\begin{bmatrix} C^{\mathsf{H}} \\ G - \lambda I \end{bmatrix} \quad = \quad \begin{bmatrix} * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \\ & * & \cdots & * \\ & & \ddots & \vdots \\ & & & * \end{bmatrix}$$

is nearly upper-triangular with $m+1$ subdiagonal lines of nonzero entries. The QR decomposition of the partial Jacobian $\mathbf{g}_{\mathbf{x}_k \cdots \mathbf{x}_1 \lambda}(G, \lambda, X)$ can then be carried out by a sequence of standard textbook Householder transformations.

9 Identifying the multiplicity support

Identification of the geometric multiplicity can be carried out with numerical rank-revealing. Let λ_0 be an initial estimate of $\lambda_* \in \mathit{eig}(A)$ in Lemma 3.1 and assume

$$|\lambda_0 - \lambda_*| \quad < \quad \theta \quad < \quad \min_{\lambda \in \mathit{eig}(A) \setminus \{\lambda_*\}} \big|\lambda - \lambda_0\big|.$$

The geometric multiplicity of λ_* can be computed as the numerical nullity of $A - \lambda_0 I$ within the error tolerance θ defined as

$$m \quad = \quad \max\big\{ j \;\big|\; \sigma_{n-j+1}(A - \lambda_0 I) < \theta \big\} \tag{18}$$

where $\sigma_i(\cdot)$ is the i-th largest singular value of $(\cdot)$. A misidentification of the geometric multiplicity can be detected during the computation. Underestimating m results in an undersized C in (1) so that both $\begin{bmatrix} A - \lambda_* I \\ C^{\mathsf{H}} \end{bmatrix}$ and the partial Jacobian $\mathbf{g}_{\lambda X}(A, \lambda_*, X_*)$ are rank-deficient. Overestimating m renders the system

$$\begin{bmatrix} A - \lambda_* I \\ C^{\mathsf{H}} \end{bmatrix} \mathbf{u} \quad = \quad \begin{bmatrix} 0 \\ T_{1:m,1} \end{bmatrix}$$

inconsistent with a large residual norm. During an iteration in which (λ_j, X_j) approaches (λ_*, X_*), a large condition number of the partial Jacobian $\mathbf{g}_{\lambda,X}(A, \lambda_j, X_j)$ indicates a likely underestimated geometric multiplicity and a large residual $\big\|\mathbf{g}(A, \lambda_j, X_j)\big\|_2$ suggests a possible overestimation.

If the geometric multiplicity is identified, it is possible to find the Segre characteristic anchor by a searching scheme based on the condition number of the Jacobian $\mathbf{g}_{\lambda X}$.

Example 1 Let

$$A = \begin{bmatrix}
0 & 4 & 0 & -4 & 0 & -2 & 1 & 0 & 0 & -1 & -1 & -1 & -1 & 2 & 1 & 0 & 0 & -1 & 0 & 0 \\
0 & 3 & 3 & -4 & 1 & 0 & 4 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & -1 & -1 \\
1 & -4 & 2 & 10 & -3 & 1 & -7 & -2 & -1 & 3 & 2 & 0 & 0 & -1 & 0 & 3 & -1 & 1 & 1 & 2 \\
-1 & -1 & 2 & 5 & -2 & -1 & -5 & -1 & -1 & 2 & 0 & -1 & 0 & 1 & 0 & 2 & -1 & -1 & -1 & 1 \\
-1 & -2 & 2 & 1 & 1 & 1 & -1 & 0 & -2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & -1 & 0 & 0 \\
1 & 4 & 1 & -12 & 4 & 2 & 13 & 3 & 0 & -4 & 0 & 0 & -2 & -1 & 1 & -6 & 1 & 1 & 0 & -3 \\
-1 & -1 & 1 & 5 & -2 & 0 & -4 & -2 & 0 & 1 & -1 & 0 & 0 & 1 & 0 & 4 & 0 & -1 & -1 & 2 \\
1 & 2 & -4 & 0 & 1 & 0 & 1 & 1 & 4 & -2 & -1 & 1 & 0 & -1 & 0 & 1 & 2 & 1 & 0 & 1 \\
0 & -5 & 2 & 10 & -5 & -1 & -10 & -2 & -1 & 6 & 3 & -2 & 0 & 0 & 0 & 3 & -3 & 0 & 1 & 2 \\
1 & 1 & 1 & -1 & 2 & 2 & 4 & 0 & 1 & -1 & -1 & 2 & 0 & -1 & 0 & 0 & 2 & 1 & -1 & 0 \\
1 & -1 & 0 & 2 & 1 & 2 & 1 & 0 & 1 & -1 & 3 & 2 & 0 & -1 & 0 & 0 & 1 & 1 & -1 & 0 \\
-1 & -3 & 0 & 5 & -1 & 2 & -4 & -1 & 0 & 1 & -1 & 4 & 4 & 1 & -2 & 2 & 0 & -1 & 0 & 1 \\
-2 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 3 & 2 & 0 & 0 & 0 & -1 & 0 & 0 \\
-3 & 4 & -1 & -4 & 0 & -2 & 1 & 0 & 0 & -1 & -1 & -1 & -2 & 5 & 2 & 0 & 0 & -1 & 1 & 0 \\
-2 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 2 & 3 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & -2 & 3 & 1 & 2 & -1 & -1 & 2 & -2 & -2 & 2 & 0 & 0 & 0 & 5 & 2 & 0 & 0 & 1 \\
6 & 3 & -6 & 3 & 6 & 4 & 7 & 0 & 7 & -7 & 1 & 5 & -2 & -6 & 1 & 0 & 8 & 6 & -1 & 0 \\
0 & 2 & -4 & -4 & 1 & -1 & 4 & 1 & 0 & -1 & 0 & -1 & -1 & 0 & 1 & -2 & 0 & 3 & 4 & -1 \\
1 & -4 & -1 & 11 & -4 & 1 & -8 & -3 & -1 & 3 & 2 & 0 & 0 & -1 & 0 & 4 & -1 & 1 & 4 & 3 \\
0 & 0 & -1 & 1 & -2 & 0 & -1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 4
\end{bmatrix}$$

with $\mathit{eig}(A) = \{2, 3\}$ of nonzero Segre characteristics $\{4, 3, 3\}$ and $\{5, 5\}$ respectively.

Applying the Francis QR algorithm implemented in Matlab yields computed eigenvalues scattered around $\lambda_1 = 2.0$ and $\lambda_2 = 3.0$:

$2.000118556521482 + 0.000118397929590i$	$3.000398490901253 + 0.001224915665189i$
$2.000118556521482 - 0.000118397929590i$	$3.000398490901253 - 0.001224915665189i$
$1.999881443477439 + 0.000118714860725i$	$3.000646066870935 + 0.000469627646058i$
$1.999881443477439 - 0.000118714860725i$	$3.000646066870935 - 0.000469627646058i$
$2.000013778528383 + 0.000018105742295i$	$3.001287762162967 + 0.000000000000000i$
$2.000013778528383 - 0.000018105742295i$	$2.999753002133234 + 0.000759191914332i$
$2.000008786021464 + 0.000020979720849i$	$2.999753002133234 - 0.000759191914332i$
$2.000008786021464 - 0.000020979720849i$	$2.998957628017279 + 0.000757681758834i$
$1.999977435451235 + 0.000002873978888i$	$2.998957628017279 - 0.000757681758834i$
$1.999977435451235 - 0.000002873978888i$	$2.999201861991639 + 0.000000000000000i$

Using two computed eigenvalues above, say

$$\tilde{\lambda}_0 = \texttt{1.999881443477439 - 0.000118714860725i} \quad \text{and} \quad \hat{\lambda}_0 = \texttt{3.001287762162967 + 0.000000000000000i}$$

as initial estimates of the defective eigenvalues, smallest singular values of $A - \tilde{\lambda}_0 I$ and $A - \hat{\lambda}_0 I$ can be computed using a rank-revealing method as

$\sigma_j(A - \tilde{\lambda}_0 I)$:	$\sigma_j(A - \hat{\lambda}_0 I)$:
...	...
0.084065699924186	0.070046163725993
0.049368630759014	0.054661269198836
0.000000000003635	0.036234932328447
0.000000000000280	**0.000000000000001**
0.000000000000001	**0.000000000000000**

indicating the geometric multiplicities 3 and 2 respectively.

Set the geometric multiplicities for the initial eigenvalue estimate $\tilde{\lambda}_0$ and $\hat{\lambda}_0$ as 3 and 2 respectively. Apply Algorithm EigenIteration with increasing input $k = 1, 2, \ldots, 6$ as estimated Segre characteristic anchors, we list the computed eigenvalues, Jacobian condition numbers and residual norms in Table 1.

At λ_1, for instance, underestimated values $k = 1, 2$ render the condition numbers of the Jacobians as large as 10^8 and the residuals to be tiny, while the overestimated value $k = 4$ leads to a drastic increase of residual from 10^{-16} to 10^{-3} but maintains the moderate Jacobian condition number, as shown in Table 1. Similar effect of increasing estimated values of the Segre characteristic anchor at λ_2 can be observed consistently. □

test	at $\lambda_1 = 2$, Segre characteristic anchor $k = 3$		
k value	computed eigenvalue	condition number	residual norm
$k = 1$	1.999881443477439 - 0.000118714860725i	560995239.6	0.000000000000001
$k = 2$	1.999999993438010 - 0.000000011324234i	147603979.2	0.000000000000001
$\rightarrow$ $k = 3$	2.000000000000000 - 0.000000000000000i	58.7	0.0000000000000006 $\leftarrow$
$k = 4$	2.109885640097783 - 0.004348977611146i	24.1	0.007

test	at $\lambda_2 = 3$, Segre characteristic anchor $k = 5$		
k value	computed eigenvalue	condition number	residual norm
$k = 1$	3.001287762162967	2161090332264.6	0.000000000000003
$k = 2$	3.001287762162967	7962600062.8	0.0000000000005
$k = 3$	3.001287762162967	4556940.4	0.000000003
$k = 4$	3.000000013572103	687859583.9	0.0000000000000007
$\rightarrow$ $k = 5$	3.000000000000000	33.9	0.0000000000000007 $\leftarrow$
$k = 6$	3.002451613695432	34.1	0.007

Table 1: Effect of increasing estimated Segre characteristic anchors: Underestimated values yield large condition numbers of the Jacobian and overestimated values lead to large residual norms. The results using the correct anchors are pointed out with arrows.

10 Improving accuracy with orthonormalization

Algorithm EigenIteration uses a simple nilpotent matrix S with only one superdiagonal line of nonzero entries. By Lemma 3.1 (iii), we can modify C and S as parameters of $\mathbf{g}$ so that the matrix component $\check{X}$ of the solution to $\mathbf{g}(A, \lambda_*, \check{X}) = \mathbf{0}$ has orthonormal columns. Assuming the parameter matrix T is set as in (3), the orthonormalization process can be carried out simply by the following process:

- Execute Algorithm EigenIteration and obtain output $\hat{\lambda}, \hat{X}, C, S$.
- Obtain the thin QR decomposition $\hat{X} = Q\,R$.
- Reset C and S with $Q^{\mathsf{H}}\,C$ and $R\,S\,R^{-1}$ respectively in the mapping $\mathbf{g}$.
- Starting from $(\lambda_0, X_0) = (\hat{\lambda}, Q)$ and execute the Gauss-Newton iteration (17) and obtain a refined eigenvalue.

The reason for such an orthonormalization is intuitively clear. When we solve for the least squares solution $(A, \tilde{\lambda}, \tilde{X})$ of the equation $\mathbf{g}(A, \lambda, X) = \mathbf{0}$ minimizing the magnitude of the residual $(A - \tilde{\lambda}\,I)\tilde{X} - \tilde{X}\,S = E$, the backward error given in Corollary 5.1 is $\|E\|_2 \left\|\tilde{X}^{\dagger}\right\|_2$. When the norm $\|\tilde{X}^{\dagger}\|_2$ is large, minimizing the residual norm $\|E\|_2$ may not achieve the highest backward accuracy. If the columns of $\tilde{X}$ are orthonormal, however, the norm $\|\tilde{X}^{\dagger}\|_2 = 1$ and the least squares solution that minimizing the residual norm $\|E\|_2$ directly minimizes the backward error bound.

Example 2 Consider the matrix

$$A = \begin{bmatrix} 2 & 1 & & & \\ & -8 & 1 & & \\ & & 2 & 1 & \\ & & & 2 & 1 \\ & -10000 & 1000 & -100 & 12 \end{bmatrix} \tag{19}$$

with an exact eigenvalue $\lambda_* = 2$ and the multiplicity support 1×5. A straightforward application of Algorithm EigenIteration in Matlab yields

$$\tilde{\lambda} = 1.999999999999748$$

$$S = \begin{bmatrix} 0 & 0.100686223197184 & 0 & 0 & 0 \\ 0 & 0 & 0.680272615629152 & 0 & 0 \\ 0 & 0 & 0 & 0.786924421181882 & 0 \\ 0 & 0 & 0 & 0 & 0.922632632948520 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\tilde{X} = \begin{bmatrix} 1.00502786434024 & 0.10210319200724 & .07627239342106 & .06542640851275 & .06584192219606 \\ -0.00000000000025 & 0.10119245986833 & .06945800549083 & .06002060904501 & .06036453955047 \\ -0.00000000000253 & 1.01192459868319 & .76341851426484 & .65486429121738 & .65902236805904 \\ -0.00000000000000 & -0.00000000000051 & .68838459356550 & .60075267245722 & .60419916522969 \\ -0.00000000000000 & -0.00000000000000 & -.00000000000052 & .54170664784191 & .55427401993992 \end{bmatrix}$$

The residual norm

$$\left\|(A - \tilde{\lambda} I)\,\tilde{X} - \tilde{X}\,S\right\|_F \approx 4.5 \times 10^{-14}$$

can not be minimized further with the unit round-off about 10^{-16} considering $\|A\|_2 \approx 10^4$. The backward error

$$\left\|(A - \tilde{\lambda} I)\,\tilde{X} - \tilde{X}\,S\right\|_F \left\|\tilde{X}^{\dagger}\right\|_2 \approx 1.3 \times 10^{-9}$$

is not small enough. After orthonormalization and resetting the resulting parameter C and S in $\mathbf{g}$ in (1), we apply the Gauss-Newton iteration again and obtain

$$\hat{\lambda} = 2.000000000000000$$

$$S = \begin{bmatrix} 0 & 0.09950371902 & -0.00990049999 & 0.00099000050 & -0.99498744208 \\ 0 & 0 & 1.00493781395 & -0.00098508732 & 0.99004950866 \\ 0 & 0 & 0 & 1.00004900870 & -0.09850873917 \\ 0 & 0 & 0 & 0 & 10050.38307728113 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\hat{X} = \begin{bmatrix} -1.0 & -0.000000000002531 & -0.000000000000000 & -0.000000000000000 & -0.000000000000000 \\ 0.0 & -0.099503719021067 & 0.009900499987341 & -0.000990000499944 & 0.994987442082474 \\ 0.0 & -0.995037190210673 & -0.000990049999192 & 0.000099000049994 & -0.099498744209908 \\ 0.0 & -0.000000000000000 & -0.999950498725976 & -0.000009900005712 & 0.009949874421338 \\ 0.0 & -0.000000000000000 & -0.000000000000000 & -0.999999505000536 & -0.000994987010945 \end{bmatrix}$$

The residual practically stays about the same magnitude

$$\left\|(A - \hat{\lambda} I)\,\hat{X} - \hat{X}\,S\right\|_F \approx 1.25 \times 10^{-14}$$

but the backward error improves substantially to

$$\left\|(A - \hat{\lambda} I)\,\hat{X} - \hat{X}\,S\right\|_F \left\|\hat{X}^{\dagger}\right\|_2 \approx 1.25 \times 10^{-14}$$

as $\|\hat{X}^{\dagger}\|_2 \approx 1$. More importantly, the forward accuracy of the computed eigenvalue improves by 3 additional accurate digits. □

When the given matrix represents perturbed data, the orthonormalization seems to be more significant in improving the accuracy, as shown in the example below.

Example 3 Using a random perturbation of magnitude about 10^{-5}, let

$$\tilde{A} = A + 10^{-5} \begin{bmatrix} -0.092 & -0.653 & -0.201 & -0.416 & -0.787 \\ -0.135 & -0.218 & 0.054 & -0.136 & -0.255 \\ 0.651 & 0.663 & -0.166 & -0.969 & -0.603 \\ -0.833 & 0.607 & 0.314 & 0.969 & -0.020 \\ -0.733 & -0.879 & 0.256 & -0.665 & -0.321 \end{bmatrix} \tag{20}$$

be the data representation of the matrix A in (19). Table 2 lists the computed eigenvalues, residual norms, backward errors and forward errors before and after orthonormalization. The results show a substantial improvement on the both forward and backward errors even though the residual magnitudes roughly stay the same. □

	before orthonormalization	after orthonormalization
computed eigenvalue	**2.00**4413315474177	**2.000000**343999377
residual norm	2.3×10^{-6}	2.9×10^{-6}
backward error	6.7×10^{-2}	2.9×10^{-6}
forward error	4.4×10^{-3}	3.4×10^{-7}

Table 2: Comparison between computing results with or without orthonormalization of the X component of the least squares solution to $\mathbf{g}(G, \lambda, X) = \mathbf{0}$ for the matrix $\tilde{A}$ in (20) at the eigenvalue $\lambda = 2$. Correct digits of computed eigenvalues are highlighted in boldface.

11 What kind of eigenvalues are ill-conditioned, and in what sense?

The well documented claim that a defective eigenvalue is infinitely sensitive to perturbations requires an oft-missing clarification: It's unbounded sensitivity is with respect to *arbitrary* perturbations. The sensitivity of a defective eigenvalue is finitely bounded by the spectral projector norm divided by the multiplicity if the perturbation is constrained to maintain the multiplicity, or by the multiplicity support condition number if the multiplicity support remains unchanged.

Furthermore, the above sensitivity assertions and clarifications are applicable on the problem of finding eigenvalues in its strictly narrow sense. In the sense of computing a multiple eigenvalue via a cluster mean provided that the cluster can be grouped correctly, the sensitivity is still bounded by spectral projector norm divided by the multiplicity. The problem of finding a defective eigenvalue in the sense of computing a pseudo-eigenvalue elaborated in this paper also enjoys a finitely bounded sensitivity in terms of the multiplicity support condition number.

Of course, the problem can still be ill-conditioned even if the sensitivity is bounded. In the following example, the matrix A has an eigenvalue of multiplicity 7 and the spectral projector norm is large, so the eigenvalue is ill-conditioned in this sense. On the other hand, the same eigenvalue is well-conditioned in multiplicity support sensitivity. Interestingly, this is not a contradiction at all. The conflicting sensitivity measures imply that the cluster mean is not accurate for approximating the eigenvalue but the pseudo-eigenvalue is, and Algorithm EigenIteration converges to the defective eigenvalue with all the digits correct.

Example 4 Let

$$A = \begin{bmatrix} 3.006 & 2 & 1.005 & -1.001 & -0.002 & -0.001 & -0.001 & -1 \\ 5 & 2 & 5 & -1 & -2 & -1 & -1 & 0 \\ -5.006 & -3 & -3.005 & 2.001 & 3.002 & 2.001 & 0.001 & 2 \\ -6 & -1 & -6 & 3 & 5 & 3 & 0 & 1 \\ -5 & -1 & -5 & 1 & 6 & 3 & 0 & 1 \\ 1 & 0 & 1 & 0 & -1 & 1 & 0 & 0 \\ -4 & -2 & -4 & 1 & 3 & 2 & 2 & 2 \\ 5 & 0 & 5 & -1 & -2 & -1 & -1 & 2 \end{bmatrix}$$

with two distinct eigenvalues in exact sense: A simple eigenvalue $\lambda_1 = 2.001$, and a defective eigenvalue $\lambda_2 = 2$ with a nonzero Segre characteristic $\{5, 2\}$ (i.e. multiplicity support is 2×2). Let P_2 be the spectral projector associated with $\lambda_2 = 2$. The defective

eigenvalue λ_2 is *both* highly ill-conditioned in spectral projector norm and almost perfectly conditioned measured by its 2×2 condition number with a sharp contrast:

$$\frac{1}{m}\|P_2\|_2 \approx 4.05 \times 10^{14} \quad \text{while} \quad \tau_{A,2\times 2}(\lambda_2) \leq 19.95.$$

This may seem to be a contradiction except it is not. Both conditions accurately measure the sensitivities of same end (finding the defective eigenvalue) through different means (cluster mean versas pseudo-eigenvalue). The Francis QR algorithm implemented in Matlab produces computed eigenvalues

```
          2.003667055821394,                  2.001912473859015 + 0.002992156370408i,
          1.996674198110247,                  2.001912473859015 - 0.002992156370408i,
          2.000000046670435,                  1.998416899175164 + 0.002994143122392i,
          1.999999953329568,                  1.998416899175164 - 0.002994143122392i.
```

There is no apparent way to group 7 computed eigenvalues to use the cluster mean for the defective eigenvalue even if we know the multiplicity is 7. Out of all 8 possible groups of 7 eigenvalues, the best approximation to $\lambda_2 = 2.0$ by the average is 2.000142850475652 with a substantial error 1.4×10^{-4} predicted by the spectral projector norm. In contrast, Algorithm EigenIteration accurately converges to $\lambda_2 = 2.0$ with an error below the unit round off 2.2×10^{-16} using the correct multiplicity support 2×2 that can easily be identified using the method described in §9, as accurately predicted by the multiplicity support condition number.

This seemingly contradicting sensitivities can be explained by the fact that there are infinitely many matrices nearby possessing a single eigenvalue of nonzero Segre characteristic $\{6, 2\}$ within 2-norm distances of 5.2×10^{-5}. Namely, such a small perturbation increases the multiplicity from 7 to 8 but can not increase the multiplicity support 2×2. Using the publicly available Matlab functionality NumericalJordanForm on the matrix A with error tolerance 10^{-5} in the software package NAClab[1] for numerical algebraic computation, we obtain approximately nearest matrix B with a single eigenvalue associated with Jordan blocks sizes 6 and 2 with first 14 digits of its entries given as

```
   3.0059955942886   1.9999978851470   1.0049959180573  -1.0010020728471  -0.0020046893569  -0.0010002300301  -0.0010132897111  -0.9999977586058
   4.9999998736434   1.9999937661529   5.0000001193777  -1.0000000065301  -2.0000000129428  -0.9999999934070  -0.9999999926252  -0.0000169379637
  -5.0060008381014  -3.0000021146845  -3.0050076499797   2.0009979267360   3.0019953102172   2.0009997699688   0.0009867094117   2.0000022421372
  -5.9999927405774  -1.0000021677309  -6.0000074892946   2.9999962015789   4.9999997701478   2.9999999999775  -0.0000002324249   0.9999978331627
  -4.9999995930006  -0.9999961877596  -5.0000095377366   0.9999880178349   5.9999870716625   3.0000002295144  -0.0000150335883   0.9999994536709
   0.9999971940837  -0.0000010574036   1.0000006545259  -0.0000047736356  -1.0000023807994   0.9999966612522  -0.0000036987224  -0.0000054161918
  -4.0000092166543  -1.9999997569827  -3.9999765043908   1.0000142841290   3.0000142782479   2.0000000005386   2.0000249853630   2.0000002431865
   4.9999998983338   0.0000026655939   5.0000001062663  -0.9999999958672  -1.9999999894366  -1.0000000065916  -1.0000000120790   2.0000133696815
```

The spectrum of B consists of a single eigenvalue $\lambda = 2.00125$. This lurking nearby matrix indicates that the multiplicity 7 of $\lambda_2 = 2.0 \in \mathit{eig}(A)$ can be increased to 8 with a small perturbation $\|A - B\|$, which is exactly the kind of cases where spectral projectors have large norms as elaborated by Kahan [9] and grouping method fails. However, those nearby defective matrices have the same multiplicity support 2×2, implying a small perturbation does not increase either the geometric multiplicity or the Segre characteristic anchor. As a result, the multiplicity support condition number is benign, and computing the defective eigenvalue via pseudo-eigenvalue is stable.

Interestingly, even though the matrix B is only known via the above empirical data, the spectral projector associated with its eigenvalue 2.00125 is known to be identity since there

[1] http://homepages.neiu.edu/∼naclab

is only one distinct eigenvalue. Consequently, the mean of all approximate eigenvalues computed by Francis QR algorithm is 2.000124999999987 with 14 digits accuracy, same as the empirical data. Algorithm EigenIteration produces the 2×2 pseudo-eigenvalue 2.000125000000078 with the same number of correct digits due to a small 2×2 condition number about 14.47.

The software NumericalJordanForm accurately produces the Jordan Canonical Forms of both matrices A and B. □

References

[1] J. V. Burke and M. L. Overton. Stable perturbations of nonsymmetric matrices. *Linear Algebra and Its Applications*, 171:249–273, 1992.

[2] F. Chaitin-Chatelin and V. Frayssé. *Lectures on Finite Precision Computations*. SIAM, Philadelphia, 1996.

[3] F. Chatelin. Ill conditioned eigenproblems. In J. Cullum and R. A. Willoughby, editors, *Large Scale Eigenvalue Problems*, North-Holland, Amsterdam, 1986. Elsevier Science Publishers B. V.

[4] J. W. Demmel. A numerical analyst's jordan canonical form. Ph.D. Dissertation, Computer Science Department, University of California, 1983.

[5] J. W. Demmel. Computing stable eigendecompositions of matrices. *Lin. Alg. and Appl.*, 79:163–193, 1986.

[6] G. H. Golub and J. H. Wilkinson. Ill-conditioned eigensystems and the computation of the Jordan canonical form. *SIAM Review*, 18:578–619, 1976.

[7] B. Kågström and A. Ruhe. Algorithm 560: JNF, an algorithm for numerical computation of the Jordan Normal Form of a complex matrix. *ACM Trans. Math. Software*, 6:437–443, 1980.

[8] B. Kågström and A. Ruhe. An algorithm for numerical computation of the Jordan normal form of a complex matrix. *ACM Trans. Math. Software*, 6:398–419, 1980.

[9] W. Kahan. Conserving confluence curbs ill-condition. Technical Report 6, Computer Science, University of California, Berkeley, 1972.

[10] T. Kato. *Perturbation Theory for Linear Operators*. Springer, Berline, Heidelberg, New York, 1966 & 1980.

[11] V. N. Kublanovskaya. On a method of solving the complete eigenvalue problem for a degenerate matrix. *USSR Computational Math. and Math. Phys.*, 6:1–14, 1968.

[12] T.-Y. Li and Z. Zeng. A rank-revealing method with updating, downdating and applications. *SIAM J. Matrix Anal. Appl.*, 26:918–946, 2005.

[13] V. Lidskii. Perturbation theory of non-conjugate operators. *U.S.S.R. Comput. Math. and Math. Phys.*, 6:73–85, 1966.

[14] R. A. Lippert and A. Edelman. The computation and sensitivity of double eigenvalues. In *Advances in computational mathematics, Lecture Notes in Pure and Appl. Math. 202*, pages 353–393, New York, 1999. Dekker.

[15] J. Moro, J. V. Burke, and M. L. Overton. On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *SIAM J. Matrix Anal. Appl.*, 18:793–817, 1997.

[16] A. Ruhe. An algorithm for numerical determination of the structure of a general matrix. *BIT*, 10:196–216, 1970.

[17] A. Ruhe. Perturbation bounds for means of eigenvalues and invariant subspaces. *BIT*, 10:343–354, 1970.

[18] S. Rump. Computational error bounds for multiple or nearly multiple eigenvalues. *Linear Algebra and its Applications*, 324:209–226, 2001.

[19] S. Rump. Eigenvalues pseudospectrum and structured perturbations. *Linear Algebra and its Applications*, 413:567–593, 2006.

[20] B. Sridhar and D. Jordan. An algorithm for calculation of the Jordan Canonical Form of a matrix. *Comput. & Elect. Engng.*, 1:239–254, 1973.

[21] L. N. Trefethen and M. Ebree. *Spectra and Pseudospectra*. Princeton University Press, Princeton and Oxford, 2005.

[22] J. H. Wilkinson. Sensitivity of eigenvalues. *Utilitas Mathematica*, 25:5–76, 1984.

[23] J. H. Wilkinson. Sensitivity of eigenvalues, II. *Utilitas Mathematica*, 30:243–286, 1986.