

Article

Observations on the Computation of Eigenvalue and Eigenvector Jacobians [†]

Andrew J. Liounis ¹, John A. Christian ^{2,*} and Shane B. Robinson ³¹ NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; andrew.j.liounis@nasa.gov² Rensselaer Polytechnic Institute, Troy, NY 12180, USA³ NASA Johnson Space Center, Houston, TX 77058, USA; shane.robinson@nasa.gov

* Correspondence: chrisj9@rpi.edu

[†] A portion of this work appears in A.J. Liounis's M.S. Thesis from West Virginia University.

Received: 24 September 2019; Accepted: 18 November 2019; Published: 20 November 2019



Abstract: Many scientific and engineering problems benefit from analytic expressions for eigenvalue and eigenvector derivatives with respect to the elements of the parent matrix. While there exists extensive literature on the calculation of these derivatives, which take the form of Jacobian matrices, there are a variety of deficiencies that have yet to be addressed—including the need for both left and right eigenvectors, limitations on the matrix structure, and issues with complex eigenvalues and eigenvectors. This work addresses these deficiencies by proposing a new analytic solution for the eigenvalue and eigenvector derivatives. The resulting analytic Jacobian matrices are numerically efficient to compute and are valid for the general complex case. It is further shown that this new general result collapses to previously known relations for the special cases of real symmetric matrices and real diagonal matrices. Finally, the new Jacobian expressions are validated using forward finite differencing and performance is compared with another technique.

Keywords: eigenvector; eigenvalue; Jacobian

1. Introduction

There are many problems that make use of the eigenvalue or eigenvector of a matrix in their solution. Because of this, it is often beneficial to be able to calculate the Jacobians of eigenvalues and eigenvectors with respect to the elements of the matrix from which they were computed. For example, eigenvalue and eigenvectors are used throughout finite-element analysis (FEA) solutions to vibration problems, where the goal is often to minimize a structure's sensitivity to various parameters through the use of eigenvalue/eigenvector derivatives [1]. As a second example, there are many instances in which the solution for the optimal estimate of a parameter vector takes the form of an eigenvalue problem—such as quaternion-based spacecraft attitude estimation [2,3] or ellipse fitting in computer vision [4]. For these types of problems, it is often important to understand the uncertainty in the optimal estimate, which requires knowledge of the eigenvector Jacobian. Other examples abound, such as direction-of-arrival of plane waves [5], epipolar geometry in computer vision [6], pose estimation in robotics [7], spacecraft optical navigation [8], or infectious disease epidemiology [9]. In all of these problems, and in many others, it can be useful (or, in some cases, necessary) to have an analytic expression for the Jacobians of eigenvalues and eigenvectors with respect to the entries in their parent matrix.

Because of the popularity of problems requiring the calculation of eigenvalues, eigenvectors, and their Jacobians, much work has already been devoted to these topics. The prior work on eigenvalue and eigenvector Jacobians can largely be grouped into three basic categories: modal expansion techniques, direct techniques, and techniques based on Nelson's method. Modal techniques mimic the idea of

modal expansion used in FEA, whereby the eigenvector Jacobian is expressed as a linear combination of the other eigenvectors of the system [10–17]. Direct solutions find the Jacobians through the solution of a linear system of equations [18–25]. Solutions based on Nelson’s method find the Jacobians by considering a homogeneous and particular solution [26–33]. Finally, there is a grouping of solutions that are not easily fit into the above categories [34–39]. For a more thorough discussion of this literature, see [39].

In light of the extensive literature already available on the topic, it may not be clear why more work is needed. After careful review, however, a few limitations are apparent with the existing methods. First, the majority of the above literature only presents solutions for the derivatives of the eigenvalues and eigenvectors with respect to a single element of the parent matrix. While these single element derivatives may be assembled to create a Jacobian, their use is cumbersome and the expressions are not compact. In the literature reviewed by these authors, expressions for the full Jacobians of the eigenvalues and eigenvectors only appear once and only for the special case of a symmetric parent matrix [34]. Second, in the majority of the literature, the derivatives are found using both the left and right eigenvectors of the system or by simultaneously solving a system of equations for both derivatives at once. While it is straightforward to find both the left and right eigensystems, calculating the left eigenvector in order to find the Jacobian of the right eigenvector is an unnecessary extra step that we seek to circumvent. Finally, much of the literature on this topic is only valid when the parent matrix and its eigenvalues and eigenvectors are real due to the choice of the normalization of the eigenvectors.

Due to these observed deficiencies, the authors of this paper proposed a new method in [39] that did not involve the left eigenvector, did not require the simultaneous solution of both the eigenvector and eigenvalue derivatives, provided the full Jacobian matrices with respect to every element of the parent matrix, and provided a solution capable of handling any matrix with real or complex elements and real or complex eigenvalues and eigenvectors. This new method found the eigenvalue Jacobian by considering the characteristic equation and solved for the eigenvector Jacobian by using the results of the eigenvalue Jacobian along with the normalization condition for the eigenvector.

While the method introduced in [39] addressed the deficiencies described above, it did produce some new issues—namely that the calculation of the eigenvalue Jacobian was computationally expensive. Additionally, the eigenvector Jacobian relied on the computation of the eigenvalue Jacobian, which could lead to numerical stability issues in large matrices ($n > 30$).

This paper introduces novel improvements to [39] that address both of these shortcomings. Specifically, the present work introduces a much simpler solution that addresses all of the original deficiencies in the eigenvalue and eigenvector derivative literature and also addresses the new issues associated with the technique in [39]. The simpler solution leads to an efficient, compact, and intuitive algorithm. The compact results presented for an arbitrary real or complex matrix (with real or complex eigenvalues and eigenvectors) are found to collapse even further for special cases and to reveal new insight into the structure of this fascinating problem. Thus, the novel approach for computing eigenvalue and eigenvector Jacobians introduced in this manuscript is exceedingly general, is applicable to a wide range of problems, and represents a significant improvement to earlier methods.

This paper is organized as follows. First, a brief discussion on the analyticity of eigenvector normalizations is presented. Then, a review of the solution proposed in [39] is presented. Next, these results are modified to arrive at the new technique, which is efficient to compute and is valid for general complex matrices. This general result is then shown to collapse to known relations from the literature for the special cases of real symmetric matrices and real diagonal matrices. The new derivatives are then validated numerically through comparison with forward finite differencing. Finally, the execution time on a digital computer is compared for the technique proposed in [39] and for the new technique proposed this paper.

2. Existence of Eigenvalue and Eigenvector Jacobians

The standard algebraic eigenvalue problem is defined as

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \tag{1}$$

where $\mathbf{A} \in \mathbb{C}^{n \times n}$ is a $n \times n$ square matrix, $\mathbf{v} \in \mathbb{C}^n$ is an eigenvector of the system, and $\lambda \in \mathbb{C}$ is an eigenvalue of the system. As should be apparent, each eigenvalue and eigenvector pair is only implicitly defined according to

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}. \tag{2}$$

It is well known that eigenvectors corresponding to simple eigenvalues are only unique when a constraint is placed on their norm (i.e., “length”). In addition, the eigenvector normalization can be expressed implicitly as

$$\mathbf{x}\mathbf{v} - \alpha = 0, \tag{3}$$

where \mathbf{x} is a row vector formed by choosing a column normalization vector and transposition method and α is a normalization constant. Although the choice of α is arbitrary, we often choose $\alpha = 1$. These two implicit vector functions can be combined to form a single implicit vector equation:

$$\mathbf{f}(\mathbf{A}, \mathbf{y}) = \begin{bmatrix} (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} \\ \mathbf{x}\mathbf{v} - \alpha \end{bmatrix} = \mathbf{0}_{n+1 \times 1}, \tag{4}$$

where $\mathbf{y} = [\mathbf{v}^T \ \lambda]^T$. To determine if eigenvalue and eigenvector derivatives exist with respect to the elements of \mathbf{A} , we must determine if both λ and \mathbf{v} can be expressed as an explicit function of \mathbf{A} . The implicit function theorem can be used to test if this is possible.

The implicit function theorem states that as long as the determinant of the Jacobian of \mathbf{f} with respect to \mathbf{y} can be found and is nonzero, then \mathbf{y} can be expressed as an explicit function of \mathbf{x} . Therefore, consider the Jacobian of \mathbf{f} as defined in Equation (4):

$$\mathbf{J} = \frac{\partial \mathbf{f}}{\partial \mathbf{y}} = \begin{bmatrix} \mathbf{A} - \lambda\mathbf{I} & -\mathbf{v} \\ \frac{\partial}{\partial \mathbf{v}}(\mathbf{x}\mathbf{v}) & 0 \end{bmatrix}, \tag{5}$$

where $\frac{\partial}{\partial \mathbf{v}}(\mathbf{x}\mathbf{v})$ will be a $n \times 1$ row vector from the rules of matrix vector calculus. Recognizing the block structure of the Jacobian, it can be shown that the determinant is given by

$$\begin{aligned} |\mathbf{J}| &= -1 |\mathbf{A} - \lambda\mathbf{I}| + \left| \mathbf{A} - \lambda\mathbf{I} + \mathbf{v} \left[\frac{\partial}{\partial \mathbf{v}}(\mathbf{x}\mathbf{v}) \right] \right| \\ &= \left| \mathbf{A} - \lambda\mathbf{I} + \mathbf{v} \left[\frac{\partial}{\partial \mathbf{v}}(\mathbf{x}\mathbf{v}) \right] \right|. \end{aligned} \tag{6}$$

Therefore, in order to evaluate the determinant of the Jacobian, we must evaluate $\partial(\mathbf{x}\mathbf{v})/\partial\mathbf{v}$ for the specific normalization described by the row vector \mathbf{x} . The following subsections discuss two potential eigenvector normalizations.

2.1. Constraining the Eigenvectors to the Unit Hypersphere

Begin with the most commonly used normalization for the standard eigenvalue/eigenvector problem, $\mathbf{x} = \mathbf{v}^H$. Substitution into Equation (3) leads to:

$$\mathbf{v}^H\mathbf{v} = \alpha \quad \rightarrow \quad -\alpha + \mathbf{v}^H\mathbf{v} = 0, \tag{7}$$

which constrains the eigenvector to lie on the hypersphere of radius α . As usual, the superscript of H indicates the Hermitian (conjugate) transpose ($\mathbf{A}^H = \text{conj}(\mathbf{A})^T$). Expressing this normalization in terms of the elements of \mathbf{v} gives

$$-\alpha + \sum_{k=1}^n \overline{v_k} v_k = -\alpha + \sum_{k=1}^n (x_k - iy_k)(x_k + iy_k) = -\alpha + \sum_{k=1}^n x_k^2 + y_k^2, \tag{8}$$

where $v_k = x_k + iy_k$ is the k^{th} element of the eigenvector \mathbf{v} and $\overline{\bullet}$ is the complex conjugate of \bullet .

We proceed by attempting to determine the partial derivative of this normalization with respect to the eigenvector in a domain around the eigenvector. First, since the eigenvector may be complex requiring a complex partial derivative, we need to check that the normalization equation is analytic in a region around the eigenvector. This is done by using the Cauchy–Riemann equations. Since the partial derivative and summation operators are linear, the required partial derivatives for checking the Cauchy–Riemann equations are

$$\frac{\partial}{\partial x_j} \text{Re} \left[\sum_{k=1}^n x_k^2 + y_k^2 \right] = 2x_j, \tag{9}$$

$$\frac{\partial}{\partial x_j} \text{Im} \left[\sum_{k=1}^n x_k^2 + y_k^2 \right] = 0, \tag{10}$$

and

$$\frac{\partial}{\partial y_j} \text{Re} \left[\sum_{k=1}^n x_k^2 + y_k^2 \right] = 2y_j, \tag{11}$$

$$\frac{\partial}{\partial y_j} \text{Im} \left[\sum_{k=1}^n x_k^2 + y_k^2 \right] = 0, \tag{12}$$

where $\text{Re}[\bullet]$ takes the real component of \bullet and $\text{Im}[\bullet]$ takes the imaginary component of \bullet . Now, using the Cauchy–Riemann equations, it is easy to see that this function is differentiable only when x_j and y_j are 0 and is analytic nowhere. In addition, this shows that this normalization is only differentiable when $\mathbf{v} = \mathbf{0}$, which is not generally a valid eigenvector. Therefore, this normalization cannot be used when the eigenvectors are complex because the normalization is not differentiable or analytic for valid eigenvectors. We do note, however, that when $\mathbf{v} \in \mathbb{R}^N$ there are no issues with using this normalization.

2.2. Constraining the Eigenvectors to a Hyperplane

Now, consider an eigenvector normalization which constrains the length of the projection of \mathbf{v} onto an arbitrary vector, \mathbf{v}_0 (so long as $\mathbf{v}_0^H \mathbf{v} \neq 0$, where we stress here that \mathbf{v}_0 is not functionally dependent on \mathbf{v} , λ , or \mathbf{A} .) Such a normalization occurs when $x = \mathbf{v}_0^H$ and constrains \mathbf{v} to lie on a hyperplane. Therefore, substituting this into Equation (3) leads to:

$$\mathbf{v}_0^H \mathbf{v} = \alpha \quad \rightarrow \quad -\alpha + \mathbf{v}_0^H \mathbf{v} = 0. \tag{13}$$

Specifically, note that the eigenvector \mathbf{v} is now constrained to lie on a hyperplane tangent to the hypersphere of radius $\alpha / \|\mathbf{v}\|$ at the point \mathbf{v}_0 . Expressing this normalization in the terms of the elements of \mathbf{v} and \mathbf{v}_0 gives

$$\begin{aligned} -\alpha + \sum_{k=1}^n \overline{v_{0k}} v_k &= -\alpha + \sum_{k=1}^n (a_k - ib_k)(x_k + iy_k) \\ &= -\alpha + \sum_{k=1}^n a_k x_k + b_k y_k + i(a_k y_k - b_k x_k), \end{aligned} \tag{14}$$

where v_k is as defined before and $v_{0k} = a_k + ib_k$ is the k^{th} element of \mathbf{v}_0 .

As before, since we are seeking a complex partial derivative, the Cauchy–Riemann equations are used to check if the partial derivative of the normalization is analytic in a region around the eigenvector. The partial derivatives with respect to the components of the j^{th} elements of the eigenvector are

$$\frac{\partial}{\partial x_j} \operatorname{Re} \left[\sum_{k=1}^n a_k x_k + b_k y_k + i(a_k y_k - b_k x_k) \right] = a_j, \tag{15}$$

$$\frac{\partial}{\partial x_j} \operatorname{Im} \left[\sum_{k=1}^n a_k x_k + b_k y_k + i(a_k y_k - b_k x_k) \right] = -b_j, \tag{16}$$

and

$$\frac{\partial}{\partial y_j} \operatorname{Re} \left[\sum_{k=1}^n a_k x_k + b_k y_k + i(a_k y_k - b_k x_k) \right] = b_j, \tag{17}$$

$$\frac{\partial}{\partial y_j} \operatorname{Im} \left[\sum_{k=1}^n a_k x_k + b_k y_k + i(a_k y_k - b_k x_k) \right] = a_j, \tag{18}$$

which satisfy the Cauchy–Riemann equations for any choice of a_j and b_j indicating that the normalization is analytic in all of \mathbb{C}^n . Since the Cauchy–Riemann equations are satisfied, the vector form of the partial derivative of the eigenvector normalization constraint with respect to the eigenvector is given as

$$\frac{\partial}{\partial \mathbf{v}} \left(\mathbf{v}_0^H \mathbf{v} - \alpha \right) = \mathbf{v}_0^H \tag{19}$$

using the rules of matrix vector calculus.

This allows the implicit function theorem to be used to check to see if the eigenvalues and eigenvectors are guaranteed to be analytic functions of \mathbf{A} on some domain centered at their nominal values. Substituting Equation (19) into Equation (6) gives

$$|\mathbf{J}| = \left| \mathbf{A} - \lambda \mathbf{I} + \mathbf{v} \mathbf{v}_0^H \right|, \tag{20}$$

which is guaranteed to be full rank and have a non-zero determinant if λ is a simple eigenvalue (defined as an eigenvalue with a multiplicity of 1). Therefore, using the implicit function theorem with this normalization, it is guaranteed that the eigenvectors and eigenvalues are analytic functions of \mathbf{A} in some domain around their nominal values for simple eigenvalues. Due to this, the rest of this paper will use the normalization $\mathbf{v}_0^H \mathbf{v} = \alpha$ and the case when the eigenvalue is simple.

3. Previous Work

In [39], the authors of this paper presented a new technique for calculating the eigenvalue and eigenvector Jacobians. This technique allows the calculation of the full Jacobian matrices for both the eigenvalues and eigenvectors using just the eigenvalue and eigenvector being considered. Some key results from this prior work are now reviewed.

Begin with the standard eigenvalue problem

$$\mathbf{A} \mathbf{v} = \lambda \mathbf{v}, \tag{21}$$

where λ is a simple eigenvalue of \mathbf{A} and \mathbf{v} is its corresponding eigenvector. Now, take the partial derivative of this relation with respect to the vectorized version of the matrix \mathbf{A} to get

$$(\mathbf{A} - \lambda \mathbf{I}) \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = \mathbf{v} \frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} - \mathbf{v}^T \otimes \mathbf{I}, \tag{22}$$

where \mathbf{A}_{vec} is formed by stacking the columns of \mathbf{A} into a column vector and \otimes indicates the Kronecker product. This results in a single equation with two unknowns, $\partial \mathbf{v} / \partial \mathbf{A}_{\text{vec}}$ and $\partial \lambda / \partial \mathbf{A}_{\text{vec}}$, and thus a second equation is needed. As described in [39], the eigenvalue derivative can be calculated by considering the characteristic equation for the matrix \mathbf{A} . The characteristic equation for \mathbf{A} can be expressed using the notation of exterior algebra [40] as

$$|\lambda \mathbf{I} - \mathbf{A}| = \sum_{k=0}^n \lambda^{n-k} (-1)^k \text{Tr} [\wedge^k \mathbf{A}] = 0, \tag{23}$$

where n is the dimension of \mathbf{A} , $|\bullet|$ indicates the determinant, $\wedge^k \mathbf{A}$ is the k^{th} exterior power of \mathbf{A} , and [41]

$$\text{Tr} [\wedge^k \mathbf{A}] = \frac{|\mathbf{Q}_k|}{k!} \tag{24}$$

$$\mathbf{Q}_k = \begin{bmatrix} \text{Tr} [\mathbf{A}] & k-1 & 0 & \dots & 0 \\ \text{Tr} [\mathbf{A}^2] & \text{Tr} [\mathbf{A}] & k-2 & \dots & 0 \\ \text{Tr} [\mathbf{A}^3] & \text{Tr} [\mathbf{A}^2] & \text{Tr} [\mathbf{A}] & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{Tr} [\mathbf{A}^k] & \text{Tr} [\mathbf{A}^{k-1}] & \text{Tr} [\mathbf{A}^{k-2}] & \dots & \text{Tr} [\mathbf{A}] \end{bmatrix}. \tag{25}$$

With the above expression for the characteristic equation, the solution for the eigenvalue Jacobian is easy to find:

$$\frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} = \frac{-\sum_{k=0}^n \lambda^{n-k} \frac{\partial c_k}{\partial \mathbf{A}_{\text{vec}}}}{\sum_{k=0}^{n-1} c_k (n-k) \lambda^{n-k-1}}, \tag{26}$$

where

$$c_k = (-1)^k \text{Tr} [\wedge^k \mathbf{A}], \tag{27}$$

$$\frac{\partial c_k}{\partial \mathbf{A}_{\text{vec}}} = \frac{(-1)^k}{k} \frac{\partial |\mathbf{Q}_k|}{\partial \mathbf{A}_{\text{vec}}}, \tag{28}$$

$$\frac{\partial |\mathbf{Q}_k|}{\partial \mathbf{A}_{\text{vec}}} = |\mathbf{Q}_k| \text{vec} \left(\mathbf{Q}_k^{-T} \right)^T \begin{bmatrix} \text{vec} (\mathbf{I})^T \\ 2 \text{vec} \left(\mathbf{A}^T \right)^T \\ 3 \text{vec} \left(\left(\mathbf{A}^2 \right)^T \right)^T \\ \vdots \\ k \text{vec} \left(\left(\mathbf{A}^{k-1} \right)^T \right)^T \\ \mathbf{0}_{1 \times n^2} \\ \text{vec} (\mathbf{I})^T \\ 2 \text{vec} \left(\mathbf{A}^T \right)^T \\ \vdots \\ (k-1) \text{vec} \left(\left(\mathbf{A}^{k-1} \right)^T \right)^T \\ \mathbf{0}_{2 \times n^2} \\ \vdots \\ \mathbf{0}_{(k-1) \times n^2} \\ \text{vec} (\mathbf{I})^T \end{bmatrix}. \tag{29}$$

Now, with the independent equation for the eigenvalue Jacobian, return to Equation (22) to solve for the eigenvector Jacobian. While it would appear simple to calculate the derivative of the eigenvectors directly from Equation (22), remember that the term $\mathbf{A} - \lambda\mathbf{I}$ is rank deficient (by the definition of an eigenvalue) and is therefore not invertible. In order to solve this problem, we must make use of a normalization condition to make the eigenvectors unique. As discussed in the preceding section (and [34]), it is important to ensure that the normalization chosen makes the eigenvector an analytic function of \mathbf{A} ; therefore, choose the normalization

$$\mathbf{v}_0^H \mathbf{v} = \alpha, \tag{30}$$

where \mathbf{v}_0 is any non-zero vector that is not orthogonal to \mathbf{v} and α is any real non-zero scalar value. In practice, it is usually chosen that numerically $\mathbf{v}_0 \equiv \mathbf{v}$, as this gives rise to a more intuitive interpretation (as we discuss later); however, even when this is the case, it is still important to remember that \mathbf{v}_0 is not a function of \mathbf{A} or \mathbf{v} .

In addition to making the eigenvectors analytic, Equation (30) also leads to another important relation. Consider the derivative of Equation (30) with respect to \mathbf{A}_{vec} ,

$$\mathbf{v}_0^H \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = \mathbf{0}_{1 \times n^2}, \tag{31}$$

which shows that the normalization vector \mathbf{v}_0 is orthogonal to the column space of the eigenvector Jacobian.

Using the above normalization condition and its derivative, perform a rank-one update to the matrix on the lefthand side of Equation (22) with the so-called *null space matrix*, \mathbf{N} , in order to make $\mathbf{A} - \lambda\mathbf{I}$ invertible. This approach is discussed further in [39]. This allows for the solution of the eigenvector derivative as

$$\frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = (\mathbf{A} - \lambda\mathbf{I} + \mathbf{N})^{-1} \left[\mathbf{v} \frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} - \mathbf{v}^T \otimes \mathbf{I} \right], \tag{32}$$

where

$$\mathbf{N} = \sigma \mathbf{v}_0 \mathbf{v}_0^H \tag{33}$$

is the Null Space Matrix and σ is a scaling term chosen to make \mathbf{N} approximately the same order as $\mathbf{A} - \lambda\mathbf{I}$ (it is usually sufficient to let the scale be $\sigma = \text{Tr}[\mathbf{A}]$). Including the Null Space Matrix, as in Equation (33), is the same as adding a zero vector due to the relation in Equation (31). Note, of course, that a similar end objective may accomplished by using the pseudoinverse of $\mathbf{A} - \lambda\mathbf{I}$ instead of including a Null Space Matrix (an example of this alternative approach may be found in [34]).

The term $\mathbf{A} - \lambda\mathbf{I}$ is rank deficient with a null space in the direction of \mathbf{v} . Additionally, the rank one Null Space Matrix only contains information in the direction of \mathbf{v}_0 . Thus, because \mathbf{v}_0 is required to not be orthogonal to \mathbf{v} , the quantity $\mathbf{A} - \lambda\mathbf{I} + \mathbf{N}$ is guaranteed to be full rank. For further discussion of the use of the Null Space Matrix and its properties, refer to [39].

This concludes the review of the technique proposed in [39] and attention is now turned to the simplification of these methods.

4. Compact Expressions for Eigenvalue and Eigenvector Jacobians

Using the insights from [39], it is now possible to arrive at a simpler and more efficient solution. Beginning again with Equation (21), left multiply by \mathbf{v}_0^H in order to form a new equation:

$$\mathbf{v}_0^H \mathbf{A} \mathbf{v} = \lambda \mathbf{v}_0^H \mathbf{v}. \tag{34}$$

Recall that, while it is common to set $\mathbf{v}_0 \equiv \mathbf{v}$, \mathbf{v}_0 is not a function of \mathbf{A} . Taking the derivative of Equation (34) leads to

$$\mathbf{v}_0^H \mathbf{A} \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} + \mathbf{v}^T \otimes \mathbf{v}_0^H = \mathbf{v}_0^H \mathbf{v} \frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} + \lambda \mathbf{v}_0^H \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} \tag{35}$$

through simple application of the chain rule and identities pertaining to the vectorization of a matrix. Note again that a superscript of T indicates a standard transpose while a superscript of H indicates the Hermitian (or conjugate) transpose. Now, recalling Equation (31), the right most term of Equation (35) vanishes. Thus, after incorporating Equation (30) and a few simple rearrangements, one finds

$$\frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} = \frac{\mathbf{v}_0^H \mathbf{A} \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} + \mathbf{v}^T \otimes \mathbf{v}_0^H}{\alpha}, \tag{36}$$

which expresses the eigenvalue Jacobian as a function of \mathbf{A} , \mathbf{v} , \mathbf{v}_0 , $\partial \mathbf{v} / \partial \mathbf{A}_{\text{vec}}$, and α .

We now turn our attention to finding a compact expression for the eigenvector Jacobian. Observe that Equations (22) and (36) create a system of two equations with two unknowns. Therefore, substituting Equation (36) into Equation (22),

$$(\mathbf{A} - \lambda \mathbf{I}) \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = \frac{\mathbf{v} \left[\mathbf{v}_0^H \mathbf{A} \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} + \mathbf{v}^T \otimes \mathbf{v}_0^H \right]}{\alpha} - \mathbf{v}^T \otimes \mathbf{I}, \tag{37}$$

which can be arranged to give

$$\left[\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v} \mathbf{v}_0^H \mathbf{A}}{\alpha} \right] \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = \frac{\mathbf{v} (\mathbf{v}^T \otimes \mathbf{v}_0^H)}{\alpha} - \mathbf{v}^T \otimes \mathbf{I} \tag{38}$$

as an equation that isolates the eigenvector derivative. This expression can be simplified to

$$\left[\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v} \mathbf{v}_0^H \mathbf{A}}{\alpha} \right] \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = \mathbf{v}^T \otimes \left(\frac{\mathbf{v} \mathbf{v}_0^H}{\alpha} - \mathbf{I} \right) \tag{39}$$

by manipulating the Kronecker products.

4.1. Eigenvector Jacobian

The objective is now to solve Equation (39) for $\partial \mathbf{v} / \partial \mathbf{A}_{\text{vec}}$. Unlike the result from [39], there is no need to incorporate a Null Space Matrix (or to use a pseudoinverse) since the term $\mathbf{A} - \lambda \mathbf{I} - \mathbf{v} \mathbf{v}_0^H \mathbf{A} / \alpha$ is already full rank and invertible as long as the eigenvalue under consideration is non-zero. This fact is straightforward to show by considering the column space of the term $\mathbf{A} - \lambda \mathbf{I}$, which will generally be rank $n - 1$ (assuming λ is simple and \mathbf{A} is full rank). Specifically, $\mathbf{A} - \lambda \mathbf{I}$ spans \mathbb{C}^{n-1} with a null space in the direction of \mathbf{v} . Now, consider the column space of the term $\mathbf{v} \mathbf{v}_0^H \mathbf{A} / \alpha$, which is rank one and spans only \mathbf{v} . Therefore, by adding $\mathbf{v} \mathbf{v}_0^H \mathbf{A} / \alpha$ to $\mathbf{A} - \lambda \mathbf{I}$, the resulting column space spans all of \mathbb{C}^n , making the overall term full rank and invertible.

In light of this fact, the solution for the eigenvector Jacobian for a non-zero eigenvalue is given by

$$\frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = \left[\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v} (\mathbf{v}_0^H \mathbf{A})}{\alpha} \right]^{-1} \left[\mathbf{v}^T \otimes \left(\frac{\mathbf{v} \mathbf{v}_0^H}{\alpha} - \mathbf{I} \right) \right], \tag{40}$$

which is a function of only \mathbf{A} , λ , \mathbf{v} , \mathbf{v}_0 , and α . Additionally, manipulating the Kronecker products allows for a final form of

$$\frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = \mathbf{v}^T \otimes \left[\left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v}\mathbf{v}_0^H \mathbf{A}}{\alpha} \right)^{-1} \left(\frac{\mathbf{v}\mathbf{v}_0^H}{\alpha} - \mathbf{I} \right) \right]. \tag{41}$$

In the case where it is necessary to find the eigenvector Jacobians when the eigenvalue is zero, the Null Space Matrix can be used as discussed in [39] to make the left-hand side invertible, resulting in

$$\frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = \mathbf{v}^T \otimes \left[\left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v}\mathbf{v}_0^H \mathbf{A}}{\alpha} + \mathbf{N} \right)^{-1} \left(\frac{\mathbf{v}\mathbf{v}_0^H}{\alpha} - \mathbf{I} \right) \right]. \tag{42}$$

4.2. Eigenvalue Jacobian

Now, consider how the eigenvector Jacobian may be used to find the eigenvalue Jacobian. Substitute the result of Equation (41) into Equation (36) and collect like terms

$$\frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} = \frac{\mathbf{v}^T}{\alpha} \otimes \left[\mathbf{v}_0^H \mathbf{A} \left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v}\mathbf{v}_0^H \mathbf{A}}{\alpha} \right)^{-1} \left(\frac{\mathbf{v}\mathbf{v}_0^H}{\alpha} - \mathbf{I} \right) + \mathbf{v}_0^H \right]. \tag{43}$$

For the zero-eigenvalue case, the equivalent expression for the eigenvalue Jacobian becomes

$$\frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} = \frac{\mathbf{v}^T}{\alpha} \otimes \left[\mathbf{v}_0^H \mathbf{A} \left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v}\mathbf{v}_0^H \mathbf{A}}{\alpha} + \mathbf{N} \right)^{-1} \left(\frac{\mathbf{v}\mathbf{v}_0^H}{\alpha} - \mathbf{I} \right) + \mathbf{v}_0^H \right]. \tag{44}$$

Focusing on the general case from Equation (43), we observe that further simplification is possible. Begin simplification by recognizing that the following are true:

$$\begin{aligned} \left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v}\mathbf{v}_0^H \mathbf{A}}{\alpha} \right) \mathbf{v} &= (\mathbf{A} - \lambda \mathbf{I}) \mathbf{v} - \left(\frac{\mathbf{v}\mathbf{v}_0^H \mathbf{A}}{\alpha} \right) \mathbf{v} \\ &= -\frac{1}{\alpha} \mathbf{v}\mathbf{v}_0^H \mathbf{A} \mathbf{v} \\ &= -\frac{\lambda}{\alpha} \mathbf{v}\mathbf{v}_0^H \mathbf{v} \\ &= -\lambda \mathbf{v} \end{aligned} \tag{45}$$

and, therefore,

$$\left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v}\mathbf{v}_0^H \mathbf{A}}{\alpha} \right)^{-1} \mathbf{v} = -\frac{1}{\lambda} \mathbf{v}. \tag{46}$$

Applying this result, the intermediate term in Equation (43) may be simplified

$$\begin{aligned} & \mathbf{v}_0^H \mathbf{A} \left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v} \mathbf{v}_0^H \mathbf{A}}{\alpha} \right)^{-1} \left(\frac{\mathbf{v} \mathbf{v}_0^H}{\alpha} - \mathbf{I} \right) \\ &= -\frac{\mathbf{v}_0^H \mathbf{A} \mathbf{v} \mathbf{v}_0^H}{\lambda \alpha} - \mathbf{v}_0^H \mathbf{A} \left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v} \mathbf{v}_0^H \mathbf{A}}{\alpha} \right)^{-1} \end{aligned} \tag{47}$$

$$\begin{aligned} &= -\frac{\mathbf{v}_0^H \mathbf{v} \mathbf{v}_0^H}{\alpha} - \mathbf{v}_0^H \mathbf{A} \left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v} \mathbf{v}_0^H \mathbf{A}}{\alpha} \right)^{-1} \\ &= -\mathbf{v}_0^H - \mathbf{v}_0^H \mathbf{A} \left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v} \mathbf{v}_0^H \mathbf{A}}{\alpha} \right)^{-1}, \end{aligned} \tag{48}$$

which, upon substitution into Equation (43), leads to

$$\frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} = \mathbf{v}^T \otimes \left[-\frac{1}{\alpha} \mathbf{v}_0^H \mathbf{A} \left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v} \mathbf{v}_0^H \mathbf{A}}{\alpha} \right)^{-1} \right], \tag{49}$$

providing a concise expression for $\partial \lambda / \partial \mathbf{A}_{\text{vec}}$ only in terms of the right eigenvector and the chosen normalization convention, $\mathbf{v}_0^H \mathbf{v} = \alpha$.

Intuitively, however, we expect the eigenvalue Jacobian to be unrelated to the choice of eigenvector normalization. Indeed, we find $\partial \lambda / \partial \mathbf{A}_{\text{vec}}$ to be the same for any choice of \mathbf{v}_0 , so long as $\mathbf{v}_0^H \mathbf{v} \neq 0$. We may show this by rewriting Equation (49) without the use of \mathbf{v}_0 or α , although doing so does require consideration of the left eigenvector. Specifically, let \mathbf{w} be the left eigenvector corresponding to the eigenvalue λ (which also has a right eigenvector \mathbf{v}),

$$\mathbf{w}^H \mathbf{A} = \lambda \mathbf{w}^H, \tag{50}$$

$$\mathbf{w}^H (\mathbf{A} - \lambda \mathbf{I}) = \mathbf{0}. \tag{51}$$

Observe, therefore, that

$$\begin{aligned} & \frac{\mathbf{w}^H}{\mathbf{w}^H \mathbf{v}} \left(\mathbf{A} - \lambda \mathbf{I} - \frac{\mathbf{v} \mathbf{v}_0^H \mathbf{A}}{\alpha} \right) \\ &= \frac{1}{\mathbf{w}^H \mathbf{v}} \left[\mathbf{w}^H (\mathbf{A} - \lambda \mathbf{I}) - \frac{\mathbf{w}^H \mathbf{v} \mathbf{v}_0^H \mathbf{A}}{\alpha} \right] \\ &= -\frac{1}{\alpha} \mathbf{v}_0^H \mathbf{A}. \end{aligned} \tag{52}$$

Substituting this result into Equation (49) produces the following compact form that is independent of the choice of eigenvector normalization

$$\frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} = \frac{\mathbf{v}^T \otimes \mathbf{w}^H}{\mathbf{w}^H \mathbf{v}}. \tag{53}$$

This result leads to a few important observations about the sensitivity of λ to perturbations in \mathbf{A} . First, this demonstrates that the eigenvalue Jacobian from Equation (49) is indeed the same for any choice of \mathbf{v}_0 other than $\mathbf{v}_0^H \mathbf{v} \approx 0$. Second, Equation (49) shows the eigenvalue Jacobian in terms of only the right eigenvector, whereas Equation (53) shows the same eigenvalue Jacobian in terms of both the right and left eigenvectors. Thus, in cases where only the right eigenvector is known (and it is not desirable to compute the corresponding left eigenvector), the result of Equation (49) provides a compact means of computing the eigenvalue Jacobian. Third, and, finally, observe the division by $\mathbf{w}^H \mathbf{v}$

on Equation (53). In general, this suggests eigenvalues will be most stable when their corresponding left and right eigenvectors are colinear (as happens in a Hermitian matrix) and become unstable as the angle between \mathbf{v} and \mathbf{w} increases ($\mathbf{w}^H \mathbf{v} \rightarrow 0$).

5. On the Choice of a Normalization Vector

As mentioned earlier, the choice of \mathbf{v}_0 is arbitrary, so long as $\mathbf{v}_0^H \mathbf{v} \neq 0$. While this is true, the choice of \mathbf{v}_0 can play an important role in the numerical stability of the system in finite precision computing. In particular, the closer the normalization vector gets to being orthogonal to the eigenvector, the more numerically unstable the system becomes. To demonstrate this phenomena consider Figure 1, which shows the condition number of the term $\mathbf{A} - \lambda \mathbf{I} - \mathbf{v} \mathbf{v}_0^H \mathbf{A} / \alpha$ normalized by the number of dimensions, n , as a function of the angle between \mathbf{v}_0 and \mathbf{v} . A higher condition number indicates a more poorly conditioned matrix prone to issues in finite precision computing.

The samples were generated by producing 10,000 random matrices for each integer value in degrees for the angle between \mathbf{v}_0 and \mathbf{v} . Figure 1 shows matrices whose elements are real values which are normally distributed with zero mean and unit variance. In these experiments, the eigenvector, \mathbf{v} , for each matrix is randomly selected. Since \mathbf{v}_0 and/or \mathbf{v} may be complex, the angle between the vectors is computed using

$$\cos \theta = \frac{\text{Re}(\mathbf{v}_0^H \mathbf{v})}{\|\mathbf{v}_0\| \|\mathbf{v}\|}. \tag{54}$$

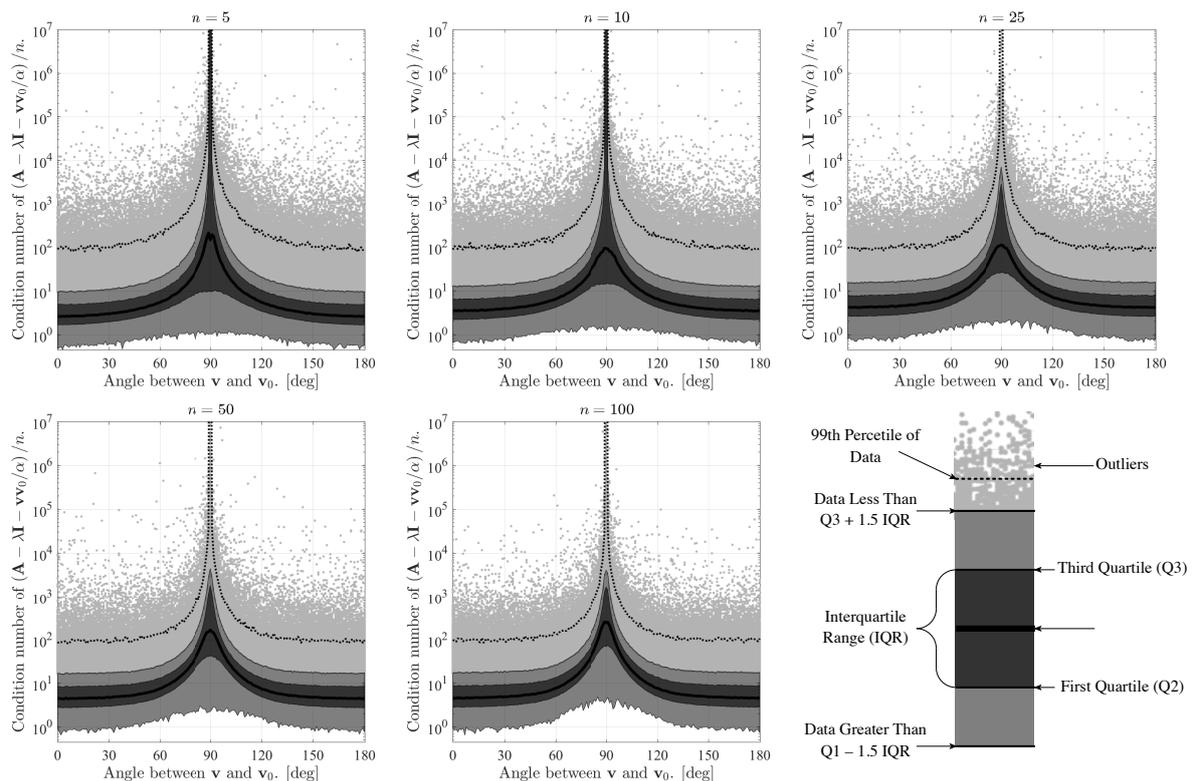


Figure 1. The condition number of the term $\mathbf{A} - \lambda \mathbf{I} - \mathbf{v} \mathbf{v}_0^H \mathbf{A} / \alpha$ normalized by the number of dimensions, n , as a function of the angle between \mathbf{v}_0 and \mathbf{v} . Here, \mathbf{A} is a real matrix. The numerical results are annotated to make the statistics more clearly visible according to the legend in the bottom right frame. Results are similar for a complex \mathbf{A} .

Figure 1 indicates that the probability of poor matrix conditioning is greatest when \mathbf{v}_0 and \mathbf{v} are nearly orthogonal. Presuming they are in random directions, as n increases, it becomes highly probable that \mathbf{v}_0 and \mathbf{v} will be nearly orthogonal.

Therefore, it is clear that choosing the normalization vector $\mathbf{v}_0 \equiv \mathbf{v}$ (leading to an angle of 0 deg; or, equivalently, $\mathbf{v}_0 \equiv -\mathbf{v}$ leading to an angle of 180 deg) will ensure the best conditioning for the system and will usually lead to the most intuitive solution. The worst conditioning occurs as $\mathbf{v}_0^H \mathbf{v} \rightarrow 0$ (angle of 90 deg). This is why $\mathbf{v}_0 \equiv \mathbf{v}$ is often chosen in practice, unless there is a compelling reason to choose otherwise.

Note that selecting $\mathbf{v}_0 = \mathbf{v}$ causes the condition number of $\mathbf{A} - \lambda \mathbf{I} - \mathbf{v}\mathbf{v}_0^H \mathbf{A} / \alpha$ to share many of the statistical properties of the condition number of \mathbf{A} . This is especially true for large values of the condition number. Edelman [42] developed probability density functions for the condition number of \mathbf{A} normalized by n when the elements of \mathbf{A} are real values with zero mean and unit variance,

$$p(x) = \frac{2x + 4}{x^3} \exp\left(-\frac{2}{x} - \frac{2}{x^2}\right). \tag{55}$$

Likewise, for the case where \mathbf{A} is composed of complex elements with real and imaginary parts which are each normally distributed with zero mean and unit variance,

$$p(x) = \frac{8}{x^3} \exp\left(-\frac{4}{x^2}\right). \tag{56}$$

Histograms for when $\mathbf{v}_0 = \mathbf{v}$ are overlaid with Edelman’s probability density functions in Figure 2. These figures show that the probability of large condition numbers occurring are well described by Edelman’s result. Unfortunately, Edelman’s density function does not work as well for small condition numbers, but these are of little consequence.

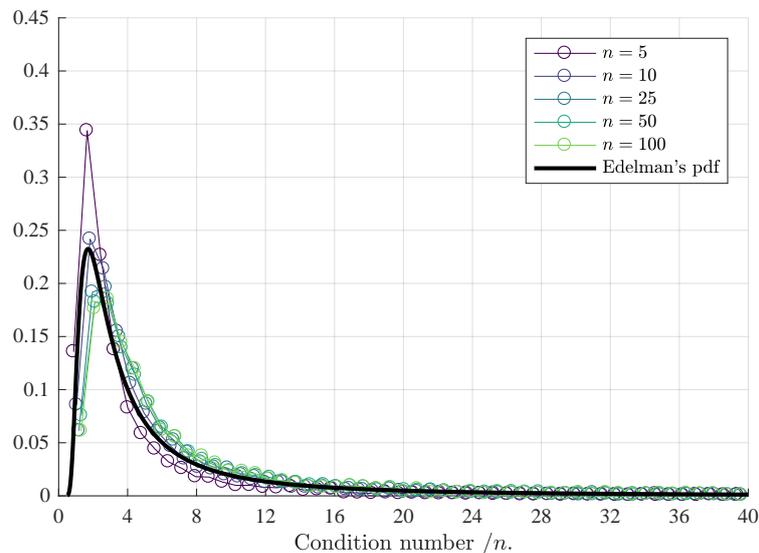


Figure 2. Histograms of the condition number of the term $\mathbf{A} - \lambda \mathbf{I} - \mathbf{v}\mathbf{v}_0^H \mathbf{A} / \alpha$ normalized by the number of dimensions, n , compared to Edelman’s probability density function for the condition number of \mathbf{A} . Here, \mathbf{A} is a real matrix.

In light of the above discussion, we stress again that \mathbf{v}_0 is not the same as \mathbf{v} . These results simply demonstrate that the best numerical performance is achieved when the values of the arbitrary vector \mathbf{v}_0 are set to be the same as \mathbf{v} . Consequently, one may wonder how the usual normalization choice of $\mathbf{v}^H \mathbf{v} = \alpha$ (instead of the $\mathbf{v}_0^H \mathbf{v} = \alpha$ normalization suggested in this paper) affects the Jacobians. Observation of the problem geometry reveals that the normalization proposed in Equation (30) ($\mathbf{v}_0^H \mathbf{v} = \alpha$) serves as a good approximation for the usual normalization of $\mathbf{v}^H \mathbf{v} = \alpha$ when \mathbf{A} is well conditioned and the values of \mathbf{v}_0 are set to that of \mathbf{v} . This is because the usual normalization constrains the eigenvectors to fall on the hypersphere of radius $\sqrt{\alpha}$, while the normalization from Equation (30)

constrains the perturbed eigenvectors to fall on the hyperplane tangent to the hypersphere of radius $\alpha / \|\mathbf{v}\|$ at \mathbf{v}_0 . Therefore, as long as the perturbation of \mathbf{A} leaves the perturbed eigenvector near the original eigenvector, the proposed normalization approximates the normalization constraining the eigenvectors to the hypersphere to first order. Furthermore, when \mathbf{v}_0 is chosen to be \mathbf{v} , the difference between the usual normalization and Equation (30) rarely needs to be considered in practice except in very rare situations, such as when attempting to numerically verify the analytic expressions as was done in the “Numerical Validation” section of this paper.

With these observations in mind, additional remarks are necessary regarding the expressions in the literature that rely on the normalization $\mathbf{v}^H \mathbf{v} = \alpha$. Recall from our earlier discussion that using $\mathbf{v}^H \mathbf{v} = \alpha$ does not result in a valid expression for the eigenvector derivatives because the normalization is not analytic. Despite this, the numerical results produced by these methods will be equivalent to the numerical results of the expressions developed in this paper when \mathbf{v}_0 is chosen to be \mathbf{v} . While the expressions for the derivatives are numerically equivalent, the derivatives from the existing literature are not valid for the normalization employed. This means that the derivatives in the existing literature cannot be used to predict the perturbations caused to eigenvectors by perturbations to \mathbf{A} when the eigenvectors are complex. The derivatives and normalization in this paper are valid for all complex eigenvectors and can be used to predict the perturbation caused to the eigenvectors by the perturbations to \mathbf{A} .

6. Simplified Cases

The goal of the work presented in this paper was to create a compact, efficient, and intuitive algorithm for the calculation of the Jacobians of an eigenvalue and eigenvector with respect to the elements of the parent matrix. Now that these expressions have been developed, it is beneficial to discuss how they simplify as assumptions are placed on the parent matrix. A variety of simplifications are possible by imposing structure on \mathbf{A} and subsequently Equations (41) and (49). This work focuses on two particularly useful simplifications as a way to gain key insights and show connections with existing literature.

6.1. Real Symmetric Parent Matrix

The first simplified case considered is when the parent matrix of the eigenvalues and eigenvectors is real and symmetric. Matrices of this structure frequently appear in practical problems from science and engineering (for instance, the Davenport solution to Wahba’s Problem [3]). In order to simplify and to parallel results from the existing literature, it is also necessary to make a choice for \mathbf{v}_0 ; therefore, choose that $\mathbf{v}_0 = \mathbf{v}$ and $\alpha = 1$. Note that, as discussed previously, this is the choice usually made in practice, as it leads to the best condition for the calculation of the eigenvector derivative.

To begin the simplifications for the symmetric case consider Equation (38), repeated here for convenience:

$$\left[\mathbf{A} - \lambda \mathbf{I} - \mathbf{v}\mathbf{v}^T \mathbf{A} \right] \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = -\mathbf{v}^T \otimes \left(\mathbf{I} - \mathbf{v}\mathbf{v}^T \right). \tag{57}$$

Note that the Hermitian transposes have been replaced by standard transposes, since the eigenvalues and eigenvectors are guaranteed to be real since \mathbf{A} is real and symmetric. In addition, note that \mathbf{v}_0 has been replaced by \mathbf{v} . Now, as discussed before, the coefficient matrix of the eigenvector Jacobian is full rank (and invertible) as long as the matrix \mathbf{A} is full rank and invertible. Making use of the fact that for an invertible matrix

$$\mathbf{A}^+ = \mathbf{A}^{-1}, \tag{58}$$

where \mathbf{A}^+ is the Moore–Penrose pseudoinverse of \mathbf{A} [43], it is possible to write

$$\frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = -\mathbf{v}^T \otimes \left[\left(\mathbf{A} - \lambda \mathbf{I} - \mathbf{v}\mathbf{v}^T \mathbf{A} \right)^+ \left(\mathbf{I} - \mathbf{v}\mathbf{v}^T \right) \right]. \tag{59}$$

From here, it is necessary to further consider the pseudoinverse term. Recognize that the pseudoinverse term in Equation (59) can be expressed as the addition of a matrix and an outer product

$$(\mathbf{A} - \lambda\mathbf{I} - \mathbf{v}\mathbf{v}^T\mathbf{A})^+ = (\mathbf{B} + \mathbf{c}\mathbf{d}^T)^+, \tag{60}$$

where $\mathbf{B} = \mathbf{A} - \lambda\mathbf{I}$, $\mathbf{c} = -\mathbf{v}$, and $\mathbf{d} = \mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Using the case i identities presented in [44] when \mathbf{A} is real and symmetric, it can be shown that

$$(\mathbf{A} - \lambda\mathbf{I} - \mathbf{v}\mathbf{v}^T\mathbf{A})^+ = (\mathbf{A} - \lambda\mathbf{I})^+ - \lambda^{-1}\mathbf{v}\mathbf{v}^T. \tag{61}$$

Substituting this result into Equation (59) yields

$$\frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = -\mathbf{v}^T \otimes \left([(\mathbf{A} - \lambda\mathbf{I})^+ - \lambda^{-1}\mathbf{v}\mathbf{v}^T] [\mathbf{I} - \mathbf{v}\mathbf{v}^T] \right). \tag{62}$$

Now, expanding the matrix multiplication in the Kronecker product gives

$$\begin{aligned} \frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = -\mathbf{v}^T \otimes & \left[(\mathbf{A} - \lambda\mathbf{I})^+ - \lambda^{-1}\mathbf{v}\mathbf{v}^T \right. \\ & \left. - (\mathbf{A} - \lambda\mathbf{I})^+ \mathbf{v}\mathbf{v}^T + \lambda^{-1}\mathbf{v}\mathbf{v}^T \mathbf{v}\mathbf{v}^T \right], \end{aligned} \tag{63}$$

which, when taking into account that for symmetric matrices the pseudoinverse has the same null space as the matrix itself, simplifies to

$$\frac{\partial \mathbf{v}}{\partial \mathbf{A}_{\text{vec}}} = -\mathbf{v}^T \otimes (\mathbf{A} - \lambda\mathbf{I})^+, \tag{64}$$

which is exactly the same result presented in [34].

With the simplified version of the eigenvector derivative in hand, the simplified eigenvalue Jacobian is trivial to find. To begin, substitute Equation (64) into Equation (36) to obtain

$$\frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} = \mathbf{v}^T \otimes \left[-\mathbf{v}^T \mathbf{A} (\mathbf{A} - \lambda\mathbf{I})^+ \right] + \mathbf{v}^T \otimes \mathbf{v}^T. \tag{65}$$

Now, making use of the fact that $\mathbf{v}^T \mathbf{A} = \lambda \mathbf{v}^T$ for symmetric matrices and the sharing of the null spaces (the original matrix and its pseudo inverse share the same null space for symmetric matrices), this reduces to

$$\frac{\partial \lambda}{\partial \mathbf{A}_{\text{vec}}} = \mathbf{v}^T \otimes \mathbf{v}^T, \tag{66}$$

which, again, is exactly the same as that presented in [34]. Note that these expressions do not assume that the parent matrix is perturbed symmetrically.

Thus, the general eigenvalue and eigenvector Jacobians presented in Equations (41) and (49) cleanly simplify to the results from [34] for the special case when \mathbf{A} is symmetric. This same result can be achieved by assuming the symmetry of \mathbf{A} and enforcing $\mathbf{v}_0 \equiv \mathbf{v}$ in Equations (22) and (36), as was done in [34].

6.2. Real Diagonal Parent Matrix

The second simplified case considered is a diagonal parent matrix with only real valued elements. In this case, the eigenvalues of the matrix are simply the diagonal elements of \mathbf{A} and the eigenvectors are the standard basis. While this case is trivial, it leads to some powerful insights into the overall problem.

6.2.1. Simplified Jacobians for a Diagonal Matrix

To develop the simplified derivatives for a diagonal matrix, begin with the simplified derivatives for the symmetric case given in Equations (64) and (66). Now, recognizing that the pseudoinverse of a diagonal matrix is just the reciprocal of the non-zero diagonal elements, the eigenvector derivative simplifies to

$$\begin{aligned} \frac{\partial \mathbf{v}_i}{\partial \mathbf{A}_{\text{vec}}} &= -\mathbf{e}_i^T \otimes (\mathbf{A} - \lambda_i \mathbf{I})^+ \\ &= - \begin{bmatrix} \mathbf{0}_{n \times (i-1)n} & (\mathbf{A} - \lambda_i \mathbf{I})^+ & \mathbf{0}_{n \times (n-i)n} \end{bmatrix}, \end{aligned} \tag{67}$$

where

$$(\mathbf{A} - \lambda_i \mathbf{I})^+ = \text{diag} \left[(\lambda_1 - \lambda_i)^{-1}, \dots, (\lambda_{i-1} - \lambda_i)^{-1}, 0, (\lambda_{i+1} - \lambda_i)^{-1}, \dots, (\lambda_n - \lambda_i)^{-1} \right] \tag{68}$$

and where \mathbf{e}_i is the i^{th} standard basis vector and the derivatives presented are for the i^{th} eigenvalue and eigenvector (at this point, it becomes necessary to distinguish the eigenvalue and eigenvector being considered). Recall that the pseudoinverse of a non-zero scalar is the reciprocal, while the pseudoinverse of a zero scalar is 0.

The eigenvalue derivative simplifies similarly:

$$\begin{aligned} \frac{\partial \lambda_i}{\partial \mathbf{A}_{\text{vec}}} &= \mathbf{e}_i^T \otimes \mathbf{e}_i \\ &= \begin{bmatrix} \mathbf{0}_{1 \times (i-1)n} & \mathbf{e}_i^T & \mathbf{0}_{1 \times (n-i)n} \end{bmatrix}. \end{aligned} \tag{69}$$

6.2.2. Perturbation to the Eigenspace of a Diagonal Matrix

With the simplified relationships in hand, it is possible to make some interesting observations on the perturbation of the eigenspace. (Again, note that it is only assumed that the parent matrix is diagonal. The perturbation matrix is not constrained to be diagonal.) The first observation is that to perturb the i^{th} eigenvalue, one must perturb the i^{th} diagonal element of the diagonal parent matrix, at least to first order. Furthermore, the perturbation to the eigenvalue in this case is exactly the perturbation to the parent matrix. While this observation should be trivial (since the diagonal elements are the eigenvalues themselves), it leads to a more interesting observation for the general case as will be discussed later. This observation can be expressed mathematically as

$$\Delta \lambda_i = \delta_i \tag{70}$$

when

$$\Delta \mathbf{A} = \delta_i \mathbf{e}_i \mathbf{e}_i^T = \delta_i \mathbf{v}_i \mathbf{v}_i^T. \tag{71}$$

The next observation is that the eigenvector is only perturbed when the i^{th} column of the parent matrix is perturbed. Furthermore, there is an analytic relationship between the change to the eigenvector, the eigenvalues, and the perturbation itself. Mathematically, this is expressed as

$$\Delta \mathbf{v}_i = - \sum_{k=1, k \neq i}^n \frac{\delta_k}{\lambda_k - \lambda_i} \mathbf{e}_k = - \sum_{k=1, k \neq i}^n \frac{\delta_k}{\lambda_k - \lambda_i} \mathbf{v}_k \tag{72}$$

when

$$\Delta \mathbf{A} = \sum_{k=1, k \neq i}^n \delta_k \mathbf{e}_k \mathbf{e}_i^T = \sum_{k=1, k \neq i}^n \delta_k \mathbf{v}_k \mathbf{v}_i^T. \tag{73}$$

These relations show that the eigenvector derivative for a diagonal matrix can be expressed as a modal expansion of the other eigenvectors if the coefficients δ_k can be calculated (which is quite simple for a diagonal matrix since the eigenvectors are the standard basis).

6.2.3. Perturbations to the Eigenspace of a Diagonalizable Matrix

Now, reconsider the case when the parent matrix is symmetric (as was done for the previous section). Since the parent matrix is symmetric, the eigenvectors will form an orthonormal basis for \mathbb{R}^n and the matrix is diagonalizable as

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \tag{74}$$

where $\mathbf{\Lambda}$ is a diagonal matrix of the eigenvalues and \mathbf{V} is an orthogonal matrix whose columns are the eigenvectors of \mathbf{A} . Substituting this into the standard eigenvalue problem gives

$$\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T\mathbf{v}_i = \lambda_i\mathbf{v}_i,$$

which can be rewritten as

$$\mathbf{\Lambda}\mathbf{V}^T\mathbf{v}_i = \lambda_i\mathbf{V}^T\mathbf{v}_i. \tag{75}$$

Since the columns of \mathbf{V} are made up of the orthogonal eigenvectors of \mathbf{A} ,

$$\mathbf{V}^T\mathbf{v}_i = \mathbf{e}_i, \tag{76}$$

which quickly reduces Equation (75) down to the diagonalized eigensystem,

$$\mathbf{\Lambda}\mathbf{e}_i = \lambda_i\mathbf{e}_i. \tag{77}$$

Now, suppose that the matrix \mathbf{A} is perturbed by adding a matrix $\Delta\mathbf{A}$. In the diagonalized space this matrix perturbation can be expressed as

$$\Delta\mathbf{\Lambda} = \mathbf{V}^T\Delta\mathbf{A}\mathbf{V}, \tag{78}$$

where $\Delta\mathbf{\Lambda}$ is an additive update to $\mathbf{\Lambda}$. The matrix $\Delta\mathbf{\Lambda}$ will not generally be diagonal.

Now that the problem has been diagonalized, the observations described above can be utilized. Since the eigenvectors in the diagonalized space form the standard basis, it is clear that the matrix $\Delta\mathbf{\Lambda}$ can be decomposed as

$$\Delta\mathbf{\Lambda} = \sum_{i=1}^n \sum_{k=1}^n \delta_{ki}\mathbf{e}_k\mathbf{e}_i^T, \tag{79}$$

where

$$\Delta\mathbf{\Lambda} = \begin{bmatrix} \delta_{11} & \delta_{12} & \dots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \dots & \delta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n1} & \delta_{n2} & \dots & \delta_{nn} \end{bmatrix} \tag{80}$$

and

$$\delta_{ki} = \mathbf{v}_k^T\Delta\mathbf{A}\mathbf{v}_i. \tag{81}$$

Now, analogous to relations in Equations (70) and (72), the perturbations in the diagonalized space are described by

$$\Delta\lambda_i = \delta_{ii}, \tag{82}$$

$$\Delta\mathbf{e}_i = - \sum_{k=1, k \neq i}^n \frac{\delta_{ki}}{\lambda_k - \lambda_i} \mathbf{e}_k. \tag{83}$$

These perturbations must now be related to perturbations in the original eigenvectors, \mathbf{v}_i . For an additive update of the diagonalized eigenvector,

$$\mathbf{v}_i + \Delta\mathbf{v}_i = \mathbf{V}(\mathbf{e}_i + \Delta\mathbf{e}_i). \tag{84}$$

Thus, it becomes apparent that the update to the original eigenvector is given by

$$\begin{aligned} \Delta\mathbf{v}_i &= \mathbf{V}\Delta\mathbf{e}_i \\ &= - \sum_{k=1, k \neq i}^n \frac{\delta_{ki}}{\lambda_k - \lambda_i} \mathbf{v}_k. \end{aligned} \tag{85}$$

Furthermore, taking the inverse of Equation (78) combined with Equation (79) gives

$$\begin{aligned} \Delta\mathbf{A} &= \mathbf{V} \sum_{i=1}^n \sum_{k=1}^n \delta_{ki} \mathbf{e}_k \mathbf{e}_i^T \mathbf{V}^T \\ &= \sum_{i=1}^n \sum_{k=1}^n \delta_{ki} \mathbf{v}_k \mathbf{v}_i^T, \end{aligned} \tag{86}$$

which shows that any perturbation can be expressed as a linear combination of the outer products of the eigenvectors of any symmetric matrix where the coefficients are found using Equation (81). While this approach is not efficient, it provides an interesting parallel to the modal expansion techniques discussed in [10–17] as well as stability theory for eigenvectors [45]. Interestingly, the result arrived at in Equation (85) is exactly that arrived at in [45] for the symmetric case.

Furthermore, if the update to \mathbf{A} is defined to be $\Delta\mathbf{A} = \partial A_{ij} \mathbf{e}_i \mathbf{e}_j^T$, then it becomes possible to find the derivatives for a perturbation to any element of \mathbf{A} as

$$\frac{\partial \mathbf{v}_i}{\partial A_{lm}} = - \sum_{k=1, k \neq i}^n \frac{v_{kl} v_{im}}{\lambda_k - \lambda_i} \mathbf{v}_k, \tag{87}$$

where $\partial \mathbf{v}_i / \partial A_{lm}$ is the partial derivative of the i^{th} eigenvector with respect to the $(l, m)^{th}$ element of the matrix \mathbf{A} , and v_{ab} is the b^{th} element of the a^{th} eigenvector of \mathbf{A} , which is very similar to what is done in [10–17]. A similar proof using left eigenvectors can be shown for the case of any diagonalizable matrix.

In summary, it once again becomes evident that the very general and very efficient expressions for eigenvalue and eigenvector Jacobians presented in this manuscript may be reduced to a variety of important special cases presented elsewhere in the literature. In addition, these simplifications provide powerful insight into the structure and dynamics of the eigenvalue and eigenvector Jacobian problem.

7. Numerical Validation

Forward finite differencing was used to validate the formulation of the new eigenvalue and eigenvector Jacobians presented in this manuscript. This provides a numerical approximation of the Jacobians which may be compared with the analytic expressions developed in this paper.

The forward finite differencing was performed by perturbing each element of the parent matrix individually in order to calculate each element of the Jacobians. The analytic derivatives from Equations (40) and (49) were then compared with the finite differences and the percent differences were calculated as

$$\% \text{ Difference} = 100 \frac{\|x_{numeric} - x_{analytic}\|}{x_{numeric}}. \tag{88}$$

This was performed for 5000 randomly generated complex matrices of size 2×2 , 5000 randomly generated complex matrices of size 3×3 , and 5000 randomly generated complex matrices of size 10×10 . The results for both the eigenvalue and eigenvector derivatives are shown in the histograms

in Figure 3. Note that, due to finite precision issues, matrices had to be ignored where the smallest component of the eigenvector derivatives was less than the perturbation size used in the finite differencing. As can be seen in the figure, the new method performed well in every instance, well below 0.1% difference for each and every element of the eigenvalue and eigenvector Jacobians. In addition, the output from the techniques derived in this paper matched to within machine precision the outputs from [39].

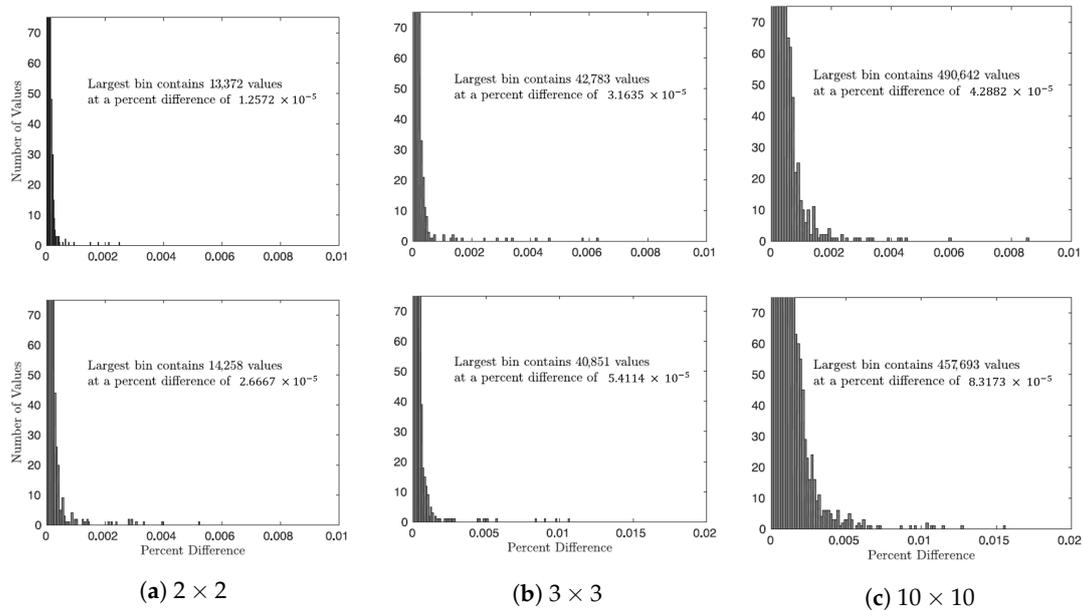


Figure 3. Histograms of percent difference between analytic derivatives computed using Equations (36) and (40) (eigenvalue derivatives top and eigenvector derivatives bottom) and finite forward differencing for 5000 randomly generated matrices of each size. The histograms are of the percent difference for each element of the eigenvalue and eigenvector derivatives (for example, for each $n \times n$ matrix there are n^2 eigenvalue derivative elements and $n \times n^2$ eigenvector derivative elements). Similar histograms are presented in [39] for the method discussed in that paper.

8. Comparison of Performance

The primary goal of the derivations presented in this paper was to decrease the computational complexity of those presented in [39]. An examination of the two formulations indicates that both techniques are $\mathcal{O}(n^4)$ due to the $n \times n$ by $n \times n^2$ multiplication in Equation (40) and the $\text{Tr}[\mathbf{A}^n]$ term in Equation (25) (assuming that the technique used to calculate the determinant is faster than $\mathcal{O}(n!)$, as this is the case in most modern linear algebra libraries). Despite the fact that both these formulations have the same upper limit on their computational complexity, it should be clear that the new formulation is much simpler, both in terms of operations performed (the formulation from [39] has two operations that are of order $\mathcal{O}(n^4)$ as opposed to one for the formulations proposed here) and in terms of memory use.

A simulation was run in an attempt to detail the increase in computational efficiency from the technique in [39] to the technique presented in this paper. The simulation was performed by applying each technique in turn to 50 randomly generated matrices (the same 50 matrices for each technique) at matrix sizes varying from 2 to 50. For each run, the computation time of each method was recorded. Finally, the minimum computation time for each matrix size was chosen for each method, and the results are shown in Figure 4. As can be seen in the figure, the new method is at minimum an order of magnitude faster and the distance between the performance of the two methods increases as the matrix size increases. In addition, note that the new method is less susceptible to numerical precision issues, as is evidenced by the cut-off of the results for the method from [39].

These simulations were performed by the authors on the campus of West Virginia University (WVU) in Morgantown, WV in 2016, using an Intel Core i7-3770 processor at 3.4 GHz, and executed within the MATLAB programming language (version R2015b).

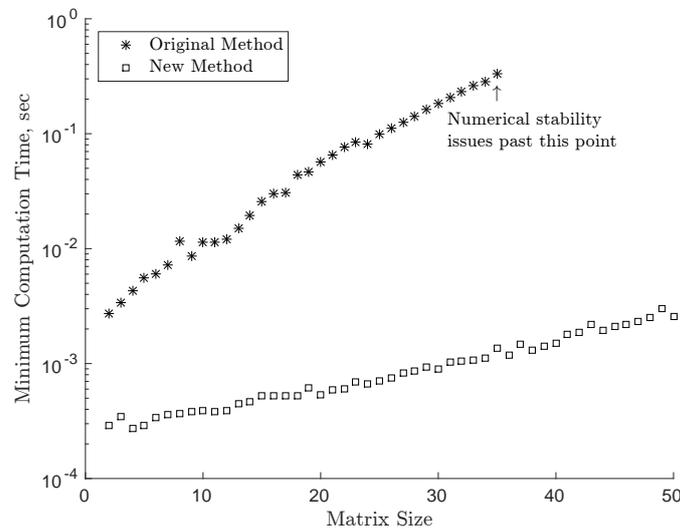


Figure 4. A plot of minimum computation time versus matrix size for the method from [39] (original method) and the method proposed in this paper (new method). Note that the method from [39] encounters numerical stability issues around a matrix size of 35 due to Equation (29). This is why there is a cut-off in the data.

9. Conclusions

A new formulation is derived for the complete Jacobians of eigenvalues and eigenvectors with respect to the elements of their parent matrix. The new solution relies on only the eigenvalue and eigenvector being considered and is valid for any unitary complex eigenvalue/eigenvector pair. Furthermore, the parent matrix may contain complex entries and need not be symmetric. As a result, the method presented here is extremely general with applications to finite-element analysis (FEA) solutions to vibration problems, fitting of an ellipse to scattered data points, quaternion-based attitude estimation, and a host of other important scientific and engineering problems.

The new eigenvalue and eigenvector Jacobians developed in this manuscript are shown to collapse to well-known results if the parent matrix is either (1) real and symmetric or (2) real and diagonal. This new method may also be reinterpreted to gain a deeper understanding of perturbations of the eigenspace.

Finally, the new eigenvalue and eigenvector Jacobians are validated by comparison with forward finite differencing. The computational performance speed of this new technique was shown to be better by a factor of 10 (or greater for large matrices) when compared with the performance of the technique proposed in [39].

Author Contributions: A.J.L.: Writing—Original Draft Preparation, Writing—Review and Editing, Formal Analysis, Conceptualization, Software, Methodology, Investigation. J.A.C.: Writing—Original Draft Preparation, Writing—Review and Editing, Formal Analysis, Conceptualization, Supervision, Project Administration. S.B.R.: Writing—Original Draft Preparation, Writing—Review and Editing, Formal Analysis, Conceptualization.

Funding: A portion of this work was made possible by NASA under award NNX13AJ25A.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Hughes, T.J. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*; Dover Publications, Inc.: Mineola, NY, USA, 2012; pp. 429–456.
- Keat, J. *Analysis of Least-Squares Attitude Determination Routine DOAOP*; Technical Report CSC/TM-77/6034; Computer Sciences Corp.: Falls Church, VA, USA, 1977.
- Markley, F.L.; Crassidis, J.L. *Fundamentals of Spacecraft Attitude Determination and Control*; Springer: New York, NY, USA, 2014; pp. 187–189.
- Fitzgibbon, A.; Pilu, M.; Fisher, R.B. Direct Least Square Fitting of Ellipses. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 476–480.
- Roy, R.; Kailath, T. ESPRIT—Estimation of Signal Parameters Via rotational Invariance Techniques. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 984–995.
- Ma, Y.; Soatto, S.; Košecá, J.; Sastry, S. *An Invitation to 3D Vision: From Images to Geometric Models*; Springer: New York, NY, USA, 2010; pp. 52–58.
- Horn, B. Closed-form Solution of Absolute Orientation Using Unit Quaternions. *J. Opt. Soc. Am.* **1987**, *4*, 629–642.
- Christian, J.A. Optical Navigation Using Planet’s Centroid and Apparent Diameter in Image. *J. Guid. Control. Dyn.* **2015**, *38*, 192–204.
- Diekmann, O.; Heesterbeek, J.; Roberts, M. The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface* **2010**, *7*, 873–885. doi:10.1098/rsif.2009.0386.
- Fox, R.; Kapoor, M. Rates of Change of Eigenvalues and Eigenvectors. *AIAA J.* **1968**, *6*, 2426–2429.
- Rogers, L.C. Derivatives of Eigenvalues and Eigenvectors. *AIAA J.* **1970**, *8*, 943–944.
- Plaut, R.; Huseyin, K. Derivatives of Eigenvalues and Eigenvectors in Non-Self-Adjoint Systems. *AIAA J.* **1973**, *11*, 250–251.
- Kalaba, R.; Spingarn, K.; Tesfatsion, L. Variational Equations for the Eigenvalues and Eigenvectors of Nonsymmetric Matrices. *J. Optim. Theory Appl.* **1981**, *33*, 1–8.
- Lim, K.; Junkins, J.; Wang, B. Re-examination of eigenvector derivatives. *J. Guid. Control. Dyn.* **1987**, *10*, 581–587.
- Andrew, A.L.; Tan, R.C. Computation of derivatives of repeated eigenvalues and corresponding eigenvectors by simultaneous iteration. *AIAA J.* **1996**, *34*, 2214–2216.
- Andrew, A.L.; Tan, R.C. Computation of Derivatives of Repeated Eigenvalues and the Corresponding Eigenvectors of Symmetric Matrix Pencils. *SIAM J. Matrix Anal. Appl.* **1998**, *20*, 78–100.
- Van Der Aa, N.; Ter Morsche, H.; Mattheij, R. Computation of Eigenvalue and Eigenvector Derivatives for a General Complex-Valued Eigensystem. *Electron. J. Linear Algebra* **2007**, *16*, 300–314.
- Garg, S. Derivatives of Eigensolutions for a General Matrix. *AIAA J.* **1973**, *11*, 1191–1194.
- Rudisill, C.S. Derivatives of Eigenvalues and Eigenvectors for a General Matrix. *AIAA J.* **1974**, *12*, 721–722.
- Rudisill, C.S.; Chu, Y.Y. Numerical Methods for Evaluating the Derivatives of Eigenvalues and Eigenvectors. *AIAA J.* **1975**, *13*, 834–837.
- Juang, J.N.; Lim, K.B. On the Eigenvalue and Eigenvector Derivatives of a General Matrix. Available online: <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19870012842.pdf> (accessed on 19 November 2019).
- Lim, K.B.; Juang, J.N.; Ghaemmaghami, P. Eigenvector derivatives of repeated eigenvalues using singular value decomposition. *J. Guid. Control. Dyn.* **1989**, *12*, 282–283.
- Juang, J.N.; Ghaemmaghami, P.; Lim, K.B. Eigenvalue and eigenvector derivatives of a nondefective matrix. *J. Guid. Control. Dyn.* **1989**, *12*, 480–486.
- Murthy, D.V.; Haftka, R.T. Derivatives of Eigenvalues and Eigenvectors of a General Complex Matrix. *Int. J. Numer. Methods Eng.* **1988**, *26*, 293–311.
- Xu, Z.; Wu, B. Derivatives of Complex Eigenvectors with Distinct and Repeated Eigenvalues. *Int. J. Numer. Methods Eng.* **2008**, *75*, 945–963.
- Nelson, R.B. Simplified Calculation of Eigenvector Derivatives. *AIAA J.* **1976**, *14*, 1201–1205.
- Cardani, C.; Mantegazza, P. Calculation of Eigenvalue and Eigenvector Derivatives for Algebraic Flutter and Divergence Eigenproblems. *AIAA J.* **1979**, *17*, 408–412.
- Ojalvo, I. *Gradients for Large Structural Models With Repeated Frequencies*; Technical Report 861789, SAE Technical Paper; University of Bridgeport: Bridgeport, CT, USA, 1986.

29. Dailey, R.L. Eigenvector Derivatives With Repeated Eigenvalues. *AIAA J.* **1989**, *27*, 486–491.
30. Mills-Curran, W.C. Calculation of Eigenvector Derivatives for Structures with Repeated Eigenvalues. *AIAA J.* **1988**, *26*, 867–871.
31. Friswell, M. The Derivatives of Repeated Eigenvalues and Their Associated Eigenvectors. *J. Vib. Acoust.* **1996**, *118*, 390–397.
32. Friswell, M.I.; Adhikari, S. Derivatives of Complex Eigenvectors Using Nelson’s Method. *AIAA J.* **2000**, *38*, 2355–2357.
33. Wu, B.; Xu, Z.; Li, Z. Improved Nelson’s Method for Computing Eigenvector Derivatives with Distinct and Repeated Eigenvalues. *AIAA J.* **2007**, *45*, 950–952.
34. Magnus, J.R. On Differentiating Eigenvalues and Eigenvectors. *Econom. Theory* **1985**, *1*, 179–191.
35. Meyer, C.D.; Stewart, G.W. Derivatives and Perturbations of Eigenvectors. *SIAM J. Numer. Anal.* **1988**, *25*, 679–691.
36. Crouzeix, M.; Philippe, B.; Sadkane, M. The Davidson Method. *SIAM J. Sci. Comput.* **1994**, *15*, 62–76.
37. Xie, H.; Dai, H. Davidson Method for Eigenpairs and Their Partial Derivatives of Generalized Eigenvalue Problems. *Commun. Numer. Methods Eng.* **2006**, *22*, 155–165.
38. De Leeuw, J. Derivatives of Generalized Eigensystems with Applications. Available online: <https://escholarship.org/content/qt2s67h3nv/qt2s67h3nv.pdf> (accessed on 19 November 2019).
39. Liounis, A.; Christian, J. Techniques for Generating Analytic Covariance Expressions for Eigenvalues and Eigenvectors. *IEEE Trans. Signal Process.* **2015**, *64*, 1808–1821. doi:10.1109/TSP.2015.2510970.
40. Winitzki, S. *Linear Algebra Via Exterior Products*; Lulu Press: Morrisville, NC, USA, 2010; pp. 69–98.
41. Bourbaki, N. *Algebra I*; Hermann: Paris, France, 1971; pp. 507–549.
42. Edelman, A. Eigenvalue and Condition Numbers of Random Matrices. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1989.
43. Penrose, R. A Generalized Inverse for Matrices. *Proc. Camb. Philos. Soc.* **1955**, *51*, 406–413.
44. Meyer, C.D. Generalized Inversion of Modified Matrices. *SIAM J. Appl. Math.* **1973**, *24*, 315–323.
45. Atkinson, K.E. *An Introduction to Numerical Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2008; pp. 591–601.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).