International Journal of Algebra, Vol. 17, 2023, no. 2, 61 - 75
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/ija.2023.91740

On Derivatives of Eigenvalues, Eigenvectors and Generalized Eigenvectors of Matrices

Namik Yener

Electrical-Electronics Engineering Department
Istanbul Ticaret University, Kucukyali E5 Kavsagi
Inonu Cad. No:4, Kucukyali, 34840 Istanbul, Turkey

Abstract

The classical linear eigenvalue problem is considered for matrices whose elements are dependent on a single real variable. Explicit expressions for derivatives of the eiegnvalues and eigenvectors are given in cases of simple and multiple eigenvalues. Recursion relations are obtained for derivatives of consecutively indexed generalized eigenvectors. Particular emphasis is placed on derivatives of eigenvectors and generalized eigenvectors to which enough coverage has not yet been provided in the present day literature.

Mathematics Subject Classification: 15A18, 15A24, 65H17

Keywords: eigenvalue; eiegnevector; generalized eigenvector; derivative

1 Introduction

In this section we do not intend to give a review of the wide body of literature that exists for the sensitivity of the eigenproblem, but we shall only attempt to identify the position of the present work within this literature.

The investigation of the sensitivity of the eigenproblem dates back to Jacobi in 1846 (see [9]). In a seminal paper in 1964 [5] Lancaster treated the problem for the classical linear eigenvalue problem

$$(A - \lambda I)x = 0 \tag{1.1}$$

when the matrix A has elements that are dependent on a single complex variable. No expressions were reported for eigenvector or generalized eigenvector derivatives, but a differentiability condition was stated for the eigenvectors. However, higher order derivatives were discussed for the eigenvalues. Kato discusses the differentiability of the eigenvalue in [4]. In [3] and [2] rates of change of eigenvalues and eigenvectors are treated when the matrix has only simple eigenvalues. In [7] the case of a matrix which is dependent on more than one parameter is taken up, but again generalized eigenvector derivatives are not touched upon.

In this work the classical linear eigenvalue problem (1.1) is considered for matrices whose elements are dependent on a single real variable. The restriction of the independent parameter to be real allows existence of derivatives more readily than in the complex variable case. On the other hand with a dependence on a single parameter complications such as having to resort to directional eigenvalue derivatives arising from the local behavior of multiple eigenvalues dependent on several parameters are not encountered [9]. We present explicit expressions for derivatives of the eigenvalues and eigenvectors in cases of both simple and multiple eigenvalues. Recursion relations are derived for derivatives of consecutively indexed generalized eigenvectors. It is to be noted that derivatives of generalized eigenvectors are generally not treated in the literature. Even though a solution to this particular problem is contained in the solution of derivatives of eigenvectors of matrix functions discussed in [1], such an approach lacks the necessary insight that can be gained by obtaining recursion relations between derivatives of consecutively indexed generalized eigenvectors. For instance by such relations a computational algorithm can be deduced to obtain these derivatives. The generalized eigenvalue problem covered in [9] considers Eq. (1.1) for A with elements that are functions of complex variables. Theorem 4.2 of this reference supplies partial derivatives for the basis vectors that span the right and left invariant subspaces of the problem corresponding to eigenvalues $\lambda_1(p), \lambda_2(p), \ldots, \lambda_r(p)$ in a neighborhood of $p = p^*$ in case of a multiple eigenvalue $\lambda_1(p^*)$ at point $p = p^*$ such that $\lambda_1(p^*) = \lambda_2(p^*) = \ldots = \lambda_r(p^*)$. However, it is noted that the vectors of these bases are usually not the eigenvectors of (1.1) associated with $\lambda_1(p), \lambda_2(p), \ldots, \lambda_r(p)$ and moreover these eigenvectors are generally not derivable at the point $p = p^*$. As we pointed out because in our treatment A has elements that are functions of a real parameter ω so that the definition of the derivative is less constricted than in the complex variable case, we can look for derivatives of eigenvectors and generalized eigenvectors in case of multiple eigenvalues. In reference [1] the general problem of derivatives of eigenvalues and eigenvectors of matrix functions is taken up. Dependence of matrix elements on more than one complex variable is assumed and partial derivatives of eigenvalues and eigenvectors are given in terms of Moore-Penrose inverse and the group inverse. Higher order derivatives are also given. Even though the derivative of a generalized eigenvector in the classical linear eigenvalue problem is a special case of the derivative of an eigenvector of a matrix function, the results given in [1] do not cover explicit expressions for this special case.

In the literature there exist alternative procedures to obtain the results of this work presented in Sections 3 through 6. However, the entirety of the work has a specific approach which is thought to be original by the author and moreover which also paves the way to determine the generalized eigenvector derivatives in Section 7. For instance the matrix F_2^s which is needed for eigenvector derivatives for multiple eigenvalues in Section 6 appears also in Section 7 for generalized eigenvector derivatives.

2 Basic Notation, Definitions and Facts

Definition 1. Eigenvalue and eigenvector of a matrix.
Let A be an n-square matrix over the field of complex numbers $\mathbb{C}$. The scalar is called $r_\nu \in \mathbb{C}$ an eigenvalue of A if there exists a nonzero column vector $f_1^{(\nu)} \in \mathbb{C}^n$ for which $Af_1^{(\nu)} = r_\nu f_1^{(\nu)}$. Every vector satisfying this relation is called an eigenvector of A belonging to the eigenvalue.

Definition 2. Generalized eigenvector of matrix A of index j with respect to an eigenvalue.
A vector $f_1^{(\nu)} \in \mathbb{C}^n$ is a generalized eigenvector of matrix A of index j (an integer ≥ 1) with respect to eigenvalue r_ν if

$$(A - r_\nu I)^p f_j^{(\nu)} = 0 \tag{2.1}$$

if and only if $p \geq j$ where p is also an integer. Note that when $j = 1$ the generalized eigenvector is termed to be a generating eigenvector. We also denote by e_ν the highest index of the generalized eigenvector of r_ν.
Therefore we now can write

$$Af_1^{(\nu)} = r_\nu f_1^{(\nu)} \tag{2.2}$$

and for the generalized eigenvectors on index greater than 1:

$$Af_1^{(\nu)} = r_\nu f_1^{(\nu)}, Af_2^{(\nu)} = r_\nu f_2^{(\nu)} + f_1^{(\nu)}, \ldots, Af_{e_\nu}^{(\nu)} = r_\nu f_{e_\nu}^{(\nu)} + f_{e_\nu - 1}^{(\nu)} \tag{2.3}$$

and additionally

$$(A - r_\nu I)^{e_\nu} f_{e_\nu}^{(\nu)} = 0. \tag{2.4}$$

Notation 2.1 Denote by B the n-square matrix whose columns are the generalized eigenvectors of A as follows:

$$B = [f_1^{(1)} \ldots f_{e_1}^1 f_1^{(2)} \ldots f_{e_2}^2 \ldots f_1^{(h)} \ldots f_{e_h}^{(h)}] \tag{2.5}$$

where $f_k^{(\mu)}$ represents the generalized eigenvector of A of index k with respect to the eigenvalue r_μ and

$$e_1 + e_2 + \cdots + e_h = n \tag{2.6}$$

Notation 2.2 If the algebraic multiplicity α_s of an eigenvalue r_s is greater than 1, each of its linearly independent eigenvectors will be denoted by $f_1^{(s_i)}$ where $1 \leq i \leq \beta_s$ and β_s is the geometric multiplicity of r_s.

Theorem 1. *The set of all the n generalized eigenvecors over $\mathbb{C}^n$ of the n-square matrix A are linearly independent.*

For a proof [6] may be referred to.

Notation 2.3 The matrix $B^{-1}\frac{dA}{d\omega}B$ where ω is a real parameter has as entry in row $(\sum_\alpha^{\mu-1} e_\alpha) + k$ and column $(\sum_\alpha^{\nu-1} e_\alpha) + j$ the value denoted by $C^\mu_{\nu jk}$ which is the coefficient of $f_k^{(\mu)}$ in the expansion $\frac{dA}{d\omega} f_j^{(\nu)} = \sum_{k,\mu} C^\mu_{\nu jk} f_k^{(\mu)}$. Or

$$\frac{dA}{d\omega} f_j^{(\nu)} = B \cdot [C^1_{\nu j1} C^1_{\nu j2} \ldots C^1_{\nu je_1} \ldots C^\mu_{\nu j1} C^\mu_{\nu j2} \ldots C^\mu_{\nu je_\mu} \ldots C^h_{\nu j1} C^h_{\nu j2} \ldots C^h_{\nu je_h}]^T \tag{2.7}$$

When ν and j run through all possible values we shall have the product of B and n columns of coefficients $C^\mu_{\nu jk}$ on the right side of (2.7), whereas the left side of (2.7) will become $\frac{dA}{d\omega}B$. Multiplying both sides by B^{-1} will yield the matrix $B^{-1}\frac{dA}{d\omega}B$ on the left and the matrix of coefficients $C^\mu_{\nu jk}$ on the right.

Theorem 2. *Number of linearly independent eigenvectors β_s with respect to an eigenvalue r_s which is a root of multiplicity α_s in the characateristic polynomial of A is equal to the number of arbitrary entries in one such linearly independent eigenvector.*

Proof. The number of linearly independent vectors β_s is equal to the nullity of $(A - r_s I)$. Also Rank of $(A - r_s I)$+nullity of $(A - r_s I) = n$. On the other hand number of arbitrary entries of an eigenvector is equal to $n-$Rank of $(A - r_s I)$. Therefore the number of arbitrary entries in an eigenvector of an eigenvalue is equal to the number of linearly independent eigenvectors corresponding to the eigenvalue. □

Notation 2.4 Take into account now rows of arbitrary entries of each such eigenvector $f_1^{(s_i)}$ corresponding to r_s which will always be at the same rows as i varies. Consider also the matrix F_1^s with columns $f_1^{(s_i)}$, the linearly independent eigenvectors of r_s. Designate the submatrix of F_1^s which has as rows the arbitrary entry rows of the linearly independent eigenvectors $f_1^{(s_i)}$ as F_2^s.

Lemma 1. *There always exists a choice of arbitrary entries of F_1^s that makes F_2^s nonsingular.*

Proof. This immediately follows from Theorem 2. For proper choice of arbitrary entries of F_1^s will make the rows of F_2^s linearly independent and the number of these rows is equal to the number of linearly independent eigenvectors β_s rendering F_2^s nonsingular. □

Lemma 2. *Generalized eigenvectors of index greater than* 1 *have as arbitrary entry rows the same arbitrary entry rows of the corresponding eigenvector (of index* 1*).*

Proof. By (2.3) we write

$$(A - r_\nu I) f_j^{(\nu)} = f_{j-1}^{(\nu)} \tag{2.8}$$

for all j with $1 \leq j \leq e_\nu$. For $j = 1$ we have the eigenvalue equation as Eq. (2.8). If n is the dimension of A, α is the rank of $(A - r_\nu I)$ and α' is the rank of the augmented matrix of (2.8) then $\alpha = \alpha' < n$. Hence we may give arbitrary values to $n - \alpha$ of the unknowns and express the other unknons in terms of these. The α unknowns which are expressed in terms of the others must be associated with some nonvanishing determinant of order α and any such set of unknowns can be used [8]. Because the coefficient matrix $(A - r_\nu I)$ remains the same for all j, the arbitrary entry row numbers in the sense of positions, of generalized eigenvectors of all indexes corresponding to an eigenvalue remain fixed. This implies that the generalized eigenvectors of higher indexes have as arbitrary entry rows the same arbitrary entry rows as the generating eigenvector of index 1. □

3 An Explicit Expression for the Derivative of an Eigenvalue of a Matrix A Which Possesses Distinct Eigenvalues

We have Eq. (2.2) for an eigenvalue r_ν and the corresponding eigenvector $f_1^{(\nu)}$ of A. Differentiating with respect to ω which A is a functtion of, we get;

$$-\frac{dr_\nu}{d\omega} f_1^{(\nu)} + \frac{dA}{d\omega} f_1^{(\nu)} = -(A - r_\nu I) \frac{df_1^{(\nu)}}{d\omega} \tag{3.1}$$

$$\frac{df_1^{(\nu)}}{d\omega} = \sum_k^n \overline{C}_{\nu k} f_1^{(k)}, \tag{3.2}$$

$$\frac{dA}{d\omega} f_1^{(\nu)} = -(A - r_\nu I)(\sum_k^n \overline{C}_{\nu k} f_1^{(k)}) + \frac{dr_\nu}{d\omega} f_1^{(\nu)}, \tag{3.3}$$

where $k \neq \nu$ in the summation sign. Or:

$$\frac{dA}{d\omega} f_1^{(\nu)} = \sum_k^n [(r_\nu - r_k)\overline{C}_{\nu k} f_1^{(k)}] + \frac{dr_\nu}{d\omega} f_1^{(\nu)}, \tag{3.4}$$

$$\frac{dA}{d\omega} f_1^{(\nu)} = B[C_{\nu 1} C_{\nu 2} \ldots C_{\nu\nu} \ldots C_{\nu n}]^T. \tag{3.5}$$

Here $C_{\nu k} = (r_\nu - r_k)\overline{C}_{\nu k}$ with $C_{\nu\nu} = \frac{dr_\nu}{d\omega}$. The coefficient matrix B is nonsingular since its columns are composed of the linearly independent eigenvectors $f_1^{(1)}, f_1^{(2)}, \ldots, f_1^{(n)}$. Hence $C_{\nu\nu} = \frac{dr_\nu}{d\omega}$ can always be solved for. It is not hard to show that $\frac{dr_\nu}{d\omega} = [B^{-1}\frac{dA}{d\omega}B]_{\nu\nu}$ where $[\cdots]_{\nu\nu}$ denotes the diagonal element of the matrix within brackets at row ν and column ν.

4 An Explicit Expression for the Derivative of a Multiple Eigenvalue of a Matrix A

Differentiating (2.4) with respect to ω yields:

$$\frac{dA}{d\omega} f_1^{(\nu)} + (A - r_\nu I)\frac{dA}{d\omega} f_2^{(\nu)} + (A - r_\nu I)^2 \frac{dA}{d\omega} f_3^{(\nu)} + \cdots + (A - r_\nu I)^{e_\nu - 2}\frac{dA}{d\omega} f_{e_\nu - 1}^{(\nu)} +$$

$$(A - r_\nu I)^{e_\nu - 1}\frac{dA}{d\omega} f_{e_\nu}^{(\nu)} - e_\nu (A - r_\nu I)^{e_\nu - 1}\frac{dr_\nu}{d\omega} f_{e_\nu}^{(\nu)} + (A - r_\nu I)^{e_\nu}\frac{df_{e_\nu}^{(\nu)}}{d\omega} = 0, \tag{4.1}$$

$$\frac{dA}{d\omega} f_1^{(\nu)} + (A - r_\nu I)\frac{dA}{d\omega} f_2^{(\nu)} + (A - r_\nu I)^2 \frac{dA}{d\omega} f_3^{(\nu)} + \cdots + (A - r_\nu I)^{e_\nu - 2}\frac{dA}{d\omega} f_{e_\nu - 1}^{(\nu)} +$$

$$(A - r_\nu I)^{e_\nu - 1}\frac{dA}{d\omega} f_{e_\nu}^{(\nu)} - e_\nu \frac{dr_\nu}{d\omega} f_1^{(\nu)} + (A - r_\nu I)^{e_\nu}\frac{df_{e_\nu}^{(\nu)}}{d\omega} + \sum_{\mu,k} \tilde{C}_{\nu k}^{\mu} f_k^{(\mu)} = 0. \tag{4.2}$$

In the above $\mu \neq \nu$ in the summation sign. Here the $\tilde{C}_{\nu k}^{\mu}$ are the new coefficients that are obtained for the expansion of $(A - r_\nu I)^{e_\nu}\frac{df_{e_\nu}^{(\nu)}}{d\omega}$. This gives:

$$\sum_{k=1}^{e_\nu} (A - r_\nu I)^{k-1}\frac{dA}{d\omega} f_k^{(\nu)} - e_\nu \frac{dr_\nu}{d\omega} f_1^{(\nu)} = -\sum_{\mu,k} \tilde{C}_{\nu k}^{\mu} f_k^{(\mu)}, \tag{4.3}$$

where in the summation sign on the right side $\mu \neq \nu$. We obtain:

$$\frac{dr_\nu}{d\omega} = \frac{1}{e_\nu a}[A_{1\gamma}A_{2\gamma}\dots A_{n\gamma}]\left[\sum_{k=1}^{e_\nu}(A - r_\nu I)^{k-1}\frac{dA}{d\omega}f_k^{(\nu)}\right], \tag{4.4}$$

where a is the determinant of B which is defined by (2.5). A_{ij} is the cofactor of the entry of B in row i and column j whereas $\gamma = (\sum_{k=1}^{\nu-1} e_k) + 1$. To arrive at the result of (4.4) the fact that B is nonsingular was utilized.

5 An Explicit Expression for the Derivative of an Eigenvector of a Matrix A Which Possesses Distinct Eigenvalues

Suppose for the unknown derivative we assume $\frac{df_1^{(\nu)}}{d\omega} = \sum_k^n \overline{C}_{\nu k} f_1^{(k)}$ as in (3.2). Then provided that $\nu \neq k$, $C_{\nu k} = (r_\nu - r_k)\overline{C}_{\nu k}$ will be the entry of $B^{-1}\frac{dA}{d\omega}B$ at column ν and row k as per Section 3. Therefore $\overline{C}_{\nu\nu}$ remains to be determined since $\overline{C}_{\nu k} = \frac{C_{\nu k}}{(r_\nu - r_k)}$ can be obtained from the proper off-diagonal entries of $B^{-1}\frac{dA}{d\omega}B$ when $\nu \neq k$. To find $\overline{C}_{\nu\nu}$ suppose one arbitrary entry of $f_1^{(\nu)}$ is in row g and denote it by $f_{1g}^{(\nu)}$. If entries of $f_1^{(k)}$ at row g are shown by $f_{1g}^{(k)}$, then $\frac{df_{1g}^{(\nu)}}{d\omega} = \sum_k^n \overline{C}_{\nu k} f_{1g}^{(k)} + \overline{C}_{\nu\nu} f_{1g}^{(\nu)}$ where in the summation $\nu \neq k$. Since a nonzero arbitrary entry $f_{1g}^{(\nu)}$ of $f_1^{(\nu)}$ will always exist for some g (otherwise $f_1^{(\nu)}$ will be a zero vector violating Definition 1), and $\overline{C}_{\nu k}$ for $\nu \neq k$ and all $f_{1g}^{(k)}$ are known $\overline{C}_{\nu\nu} = \frac{1}{f_{1g}^{(\nu)}}\left(\frac{df_{1g}^{(\nu)}}{d\omega} - \sum_k^n \overline{C}_{\nu k} f_{1g}^{(k)}\right)$, where in the summation $\nu \neq k$. In other words $\overline{C}_{\nu k}$ can all be found from the entries of the matrix $B^{-1}\frac{dA}{d\omega}B$. More concretely the explicit expression for $\frac{df_1^{(\nu)}}{d\omega}$ is:

$$\frac{df_1^{(\nu)}}{d\omega} = \sum_{\substack{k \\ k\neq\nu}}^{n} \frac{C_{\nu k}}{(r_\nu - r_k)} f_1^{(k)} + \frac{1}{f_{1g}^{(\nu)}}\left[\frac{df_{1g}^{(\nu)}}{d\omega} - \sum_{\substack{k \\ k\neq\nu}}^{n} \frac{C_{\nu k}}{(r_\nu - r_k)} f_{1g}^{(k)}\right] f_1^{(\nu)}. \tag{5.1}$$

6 An Explicit Expression for the Derivative of an Eigenvector of a Multiple Eigenvalue

Again for an eigenvector $f_1^{(\nu)}$

$$-\frac{dr_\nu}{d\omega} f_1^{(\nu)} + \frac{dA}{d\omega} f_1^{(\nu)} = -(A - r_\nu I)\frac{df_1^{(\nu)}}{d\omega} \tag{6.1}$$

holds as in Eqution (3.1). Set $\frac{df_1^{(\nu)}}{d\omega} = \sum_{k,\mu} \overline{C}^{\mu}_{\nu 1k} f_k^{(\mu)}$. Our task is to determine the $\overline{C}^{\mu}_{\nu 1k}$ so that $\frac{df_1^{(\nu)}}{d\omega}$ can be found. Now note that

$$(A - r_\nu I)\overline{C}^{\mu}_{\nu 1k} f_k^{(\mu)} = \overline{C}^{\mu}_{\nu 1k} \left[(r_\mu - r_\nu) f_k^{(\mu)} + f_{k-1}^{(\mu)} \right] \tag{6.2}$$

Recalling Eq. (2.7) for $\frac{dA}{d\omega} f_j^{(\nu)}$, from (6.1) and (6.2) we can write

$$\sum_{k,\mu} C^{\mu}_{\nu 1k} f_k^{(\mu)} - \frac{dr_\nu}{d\omega} f_1^{(\nu)} = -\sum_{k,\mu} \overline{C}^{\mu}_{\nu 1k} \left[(r_\mu - r_\nu) f_k^{(\mu)} + f_{k-1}^{(\mu)} \right] \tag{6.3}$$

We shall outline the procedure that will yield the coefficients $\overline{C}^{\mu}_{\nu 1k}$ in the following two subsections.

6.1 $r_\mu \neq r_\nu$ case

Possible conditions are: 1) Both r_μ and r_ν have more than one linearly independent eigenvector. 2) Neither has more than one linearly independent eigenvector. 3) Only one of r_μ or r_ν has more than one linearly independent eigenvector. For all these three conditions, from (6.1) and (6.3) and by noting the linear independence of the $f_k^{(\mu)}$ one can get

$$C^{\mu}_{\nu 1k} = -\overline{C}^{\mu}_{\nu 1k}(r_\mu - r_\nu) - \overline{C}^{\mu}_{\nu 1(k+1)} \tag{6.4}$$

when $k = 1, 2, \ldots, e_\mu - 1$. For $k = e_\mu$ (6.4) takes the form $C^{\mu}_{\nu 1 e_\mu} = -\overline{C}^{\mu}_{\nu 1 e_\mu}(r_\mu - r_\nu)$. The following formula can be obtained from the set of equations thus generated:

$$\overline{C}^{\mu}_{\nu 1k} = \sum_{q=1}^{e_\mu - k + 1} (-1)^q \frac{C^{\mu}_{\nu 1(k-1+q)}}{(r_\mu - r_\nu)^q} \tag{6.5}$$

As noted above in Notation 2.3 $C^{\mu}_{\nu 1k}$ in this summation is the entry of $B^{-1}\frac{dA}{d\omega}B$ in row $(\sum_\alpha^{\mu-1} e_\alpha) + k$ and column $(\sum_\alpha^{\nu-1} e_\alpha) + 1$.

6.2 $r_\mu = r_\nu$ case

Solution of the problem is given according to the below possible conditions: 1) r_ν is not a multiple eigenvalue or r_ν is a multiple eigenvalue with not more than one linearly independent eigenvector. 2) r_ν is a multiple eigenvlaue with more than one linearly independent eigenvector.

6.2.1 r_ν has only one linearly independent eigenvector

For this first condition (6.1) and (6.3) will yield for $k = 1$ in (6.3);

$$\overline{C}^{\nu}_{\nu 12} = -C^{\nu}_{\nu 11} + \frac{dr_\nu}{d\omega} \tag{6.6}$$

When $k > 1$ in (6.3):

$$\overline{C}^{\nu}_{\nu 1k} = -C^{\nu}_{\nu 1(k-1)}, \tag{6.7}$$

will be true where $C^{\nu}_{\nu 1k)}$ are again entries of $B^{-1}\frac{dA}{d\omega}B$ and completely yield coefficients in $\frac{df_1^{(\nu)}}{d\omega} = \sum_{k,\mu} \overline{C}^{\mu}_{\nu 1k} f_k^{(\mu)}$ for this condition except $\overline{C}^{\nu}_{\nu 11}$. To find $\overline{C}^{\nu}_{\nu 11}$ suppose the arbitrary entries of $f_k^{(\mu)}$ at row g are shown by $f_{kg}^{(\mu)}$. Then for this row g:

$$\frac{df_{1g}^{(\nu)}}{d\omega} = \sum_{k,\mu} \overline{C}^{\mu}_{\nu 1k} f_{kg}^{(\mu)} = \left[\sum_{k,\mu} \overline{C}^{\mu}_{\nu 1k} f_{kg}^{(\mu)}\right] + \overline{C}^{\nu}_{\nu 11} f_{1g}^{(\nu)}, \tag{6.8}$$

hence

$$\overline{C}^{\nu}_{\nu 11} = \frac{1}{f_{1g}^{(\nu)}} \left[\frac{df_{1g}^{(\nu)}}{d\omega} - \sum_{k,\mu} \overline{C}^{\mu}_{\nu 1k} f_{kg}^{(\mu)}\right]. \tag{6.9}$$

where for the summation within brackets in (6.8) and the summation in (6.9) $k = 1$ and $\mu = \nu$ not at the same time. Thus observing (6.5) as well

$$\frac{df_1^{(\nu)}}{d\omega} = \sum_{\substack{k,\mu \\ \mu\neq\nu}} \sum_{q=1}^{e_\mu - k+1} (-1)^q \frac{C^{\mu}_{\nu 1(k-1+q)}}{(r_\mu - r_\nu)^q} \left[f_k^{(\mu)} - \frac{f_{kg}^{(\mu)}}{f_{1g}^{(\nu)}} f_1^{(\nu)}\right] + \frac{1}{f_{1g}^{(\nu)}} \frac{df_{1g}^{(\nu)}}{d\omega} f_1^{(\nu)} +$$

$$\left(-C^{\nu}_{\nu 11} + \frac{dr_\nu}{d\omega}\right) f_2^{(\nu)} - \sum_{k=3}^{e_\nu} C^{\nu}_{\nu 1(k-1)} f_k^{(\nu)}. \tag{6.10}$$

6.2.2 r_ν has more than one linearly independent eigenvector

Suppose that in (6.3) μ and ν correspond to the generating eigenvectors $f_1^{(s_t)}$ and $f_1^{(s_p)}$ of the same eigenvalue r_s, because $r_\mu = r_\nu = r_s$. In this way we have r_ν with more than one linearly independent eigenvector. Eq. (6.3) for $k > 1$ will yield:

$$\overline{C}^{s_t}_{(s_p)1(k+1)} = -C^{s_t}_{(s_p)1k}, \tag{6.11}$$

and for $k = 1$:

$$\overline{C}^{s_t}_{(s_p)12} = -C^{s_t}_{(s_p)11}. \tag{6.12}$$

Suppose in (6.3) μ and ν correspond to the same generating eigenvector $f_1^{(s_i)}$. (6.6) and (6.7) will then be replaced by (6.13) and (6.14) below. For $k=1$ in (6.3):

$$\overline{C}^{s_i}_{(s_i)12} = -C^{s_i}_{(s_i)11} + \frac{dr_\nu}{d\omega}, \tag{6.13}$$

and for $k>1$:

$$\overline{C}^{s_i}_{(s_i)1(k+1)} = -C^{s_i}_{(s_i)1k}. \tag{6.14}$$

But $\overline{C}^{s_t}_{(s_p)11}$ and $\overline{C}^{s_i}_{(s_i)11}$ cannot be obtained from (6.11) through (6.14). These coefficients are needed in the expansion of $\frac{df_1^{(\nu)}}{d\omega}$ in terms of generalized eigenvectors, because:

$$\frac{df_1^{(s_p)}}{d\omega} = \left[\sum_{k,\mu} \overline{C}^{\mu}_{(s_p)1k} f_k^{(\mu)}\right] + \overline{C}^{s_1}_{(s_p)11} f_1^{(s_1)} + \overline{C}^{s_2}_{(s_p)11} f_1^{(s_2)} + \cdots + \overline{C}^{s_{\beta_s}}_{(s_p)11} f_1^{(s_{\beta_s})}. \tag{6.15}$$

The unknowns $\overline{C}^{s_t}_{(s_p)11}$ and $\overline{C}^{s_i}_{(s_i)11}$ are shown outside the summation sign in (6.15). Here β_s is the geometric multiplicity of r_s. For an arbitrary entry row g:

$$\frac{df_{1g}^{(s_p)}}{d\omega} = \left[\sum_{k,\mu} \overline{C}^{\mu}_{(s_p)1k} f_{kg}^{(\mu)}\right] + \overline{C}^{s_1}_{(s_p)11} f_{1g}^{(s_1)} + \overline{C}^{s_2}_{(s_p)11} f_{1g}^{(s_2)} + \cdots + \overline{C}^{s_{\beta_s}}_{(s_p)11} f_{1g}^{(s_{\beta_s})}. \tag{6.16}$$

In the summation signs in (6.15) and (6.16) never $\mu = s_p (1 \le p \le \beta_s)$ and $k=1$ at the same time. When g runs through all possible values we shall have a system of equations to replace (6.16) for the unknowns $\overline{C}^{s_1}_{(s_p)11}, \overline{C}^{s_2}_{(s_p)11}, \overline{C}^{s_3}_{(s_p)11}, \ldots, \overline{C}^{s_{\beta_s}}_{(s_p)11}$ whose coefficient matrix is F_2^s is as defined in Notation 2.4 of Section 2. Once these unknowns are found by solving (6.16), observing (6.5) as well we can write

$$\frac{df_1^{(s_p)}}{d\omega} = \sum_{\substack{k,\mu \\ \mu\neq\nu}} \sum_{q=1}^{e_\mu-k+1} (-1)^q \frac{C^{\mu}_{\nu 1(k-1+q)}}{(r_\mu - r_\nu)^q} f_k^{(\mu)} + \sum_{s_t=s_1}^{s_{\beta_s}} \overline{C}^{s_t}_{(s_p)11} f_1^{(s_t)} - \sum_{\substack{s_t=s_1 \\ s_t\neq s_p}}^{s_{\beta_s}} C^{s_t}_{(s_p)11} f_2^{(s_t)} +$$
$$\left[-C^{s_p}_{(s_p)11} + \frac{dr_{s_p}}{d\omega}\right] f_2^{(s_p)} - \sum_{s_t=s_1}^{s_{\beta_s}} \sum_{k=3}^{e_{s_t}} C^{s_t}_{(s_p)1(k-1)} f_k^{(s_t)}. \tag{6.17}$$

for the derivative $\frac{df_1^{(s_p)}}{d\omega}$ of one of the linearly independent eigenvectors corresponding to r_s. The first double summation term on the right is for the $r_\mu \neq r_\nu$ condition. The second term is to be computed using the solution set of Eq. (6.16) for the unkowns $\overline{C}^{s_1}_{(s_p)11}, \overline{C}^{s_2}_{(s_p)11}, \overline{C}^{s_3}_{(s_p)11}, \ldots, \overline{C}^{s_{\beta_s}}_{(s_p)11}$. The third term is for expansion terms with generalized eigenvectors $f_k^{(s_t)}$ of r_s different than $f_2^{(s_p)}$ when $k=2$. The fourth term is the same for when μ and ν correspond to the same generating eigenvector $f_1^{(s_p)}$ with $k=2$. The last term shown with a double sum is for when $k \geq 3$.

7 Derivatives of Generalized Eigenvectors of a Defective Matrix A by Recursion Formulae

The $\overline{C}^{\mu}_{\nu 1k}$ determined in Section 6 yield $\frac{df_1^{(\nu)}}{d\omega}$. The ideas developed in this section yield $\frac{df_j^{(\nu)}}{d\omega}$ when $\frac{df_{j-1}^{(\nu)}}{d\omega}$ has been determined. In this way it is possible to find derivatives of generalized eigenvectors of all indexes. From relations in (2.3) by differentiating when $j \geq 2$ and substituting according to $\frac{df_j^{(\nu)}}{d\omega} = \sum_{k,\mu} \overline{C}^{\mu}_{\nu jk} f_k^{(\mu)}$ and $\frac{df_{(j-1)}^{(\nu)}}{d\omega} = \sum_{k,\mu} \overline{C}^{\mu}_{\nu(j-1)k} f_k^{(\mu)}$ one can get:

$$-\frac{dr_\nu}{d\omega} f_j^{(\nu)} + \frac{dA}{d\omega} f_j^{(\nu)} = -(A - r_\nu I)\left[\sum_{k,\mu} \overline{C}^{\mu}_{\nu jk} f_k^{(\mu)}\right] + \sum_{k,\mu} \overline{C}^{\mu}_{\nu(j-1)k} f_k^{(\mu)}. \quad (7.1)$$

Recursion formulae relating $\overline{C}^{\mu}_{\nu jk}$ and $\overline{C}^{\mu}_{\nu(j-1)k}$ are given in two subsections: $r_\mu \neq r_\nu$ case in 7.1 and $r_\mu = r_\nu$ case (recursion formulae relating $\overline{C}^{\nu}_{\nu jk}$ and $\overline{C}^{\nu}_{\nu(j-1)k}$) in 7.2.

7.1 $r_\mu \neq r_\nu$ case

Possible conditions are: 1) Both r_μ and r_ν have more than one linearly independent eigenvector. 2) Neither has more than one linearly independent eigenvector. 3) Only one of r_μ or r_ν has more than one linearly independent eigenvector. For all these three conditions, from (7.1) and series expansion of $\frac{dA}{d\omega} f_j^{(\nu)}$ in terms of $f_k^{(\mu)}$ as $\frac{dA}{d\omega} f_j^{(\nu)} = \sum_{k,\mu} C^{\mu}_{\nu jk} f_k^{(\mu)}$, the following can be obtained:

$$C^{\mu}_{\nu j e_\mu} = -\overline{C}^{\mu}_{\nu j e_\mu}(r_\mu - r_\nu) + \overline{C}^{\mu}_{\nu(j-1)e_\mu}$$

$$C^{\mu}_{\nu j(e_\mu-1)} = -\overline{C}^{\mu}_{\nu j e_\mu} - \overline{C}^{\mu}_{\nu j(e_\mu-1)}(r_\mu - r_\nu) + \overline{C}^{\mu}_{\nu(j-1)(e_\mu-1)}$$

....

$$C^{\mu}_{\nu j1} = -\overline{C}^{\mu}_{\nu j2} - \overline{C}^{\mu}_{\nu j1}(r_\mu - r_\nu) + \overline{C}^{\mu}_{\nu(j-1)1} \quad (7.2)$$

From the above set we obtain:

$$\overline{C}^{\mu}_{\nu jk} = \sum_{q=1}^{e_\mu - k+1} (-1)^q \frac{C^{\mu}_{\nu j(k-1+q)} - \overline{C}^{\mu}_{\nu(j-1)(k-1+q)}}{(r_\mu - r_\nu)^q}. \quad (7.3)$$

This way for the case $r_\mu \neq r_\nu$ we have found all the coefficients $\overline{C}^{\mu}_{\nu jk}$ in $\frac{df_j^{(\nu)}}{d\omega} = \sum_{k,\mu} \overline{C}^{\mu}_{\nu jk} f_k^{(\mu)}$ in terms of $\overline{C}^{\mu}_{\nu(j-1)k}$.

7.2 $r_\mu = r_\nu$ case

Solution of the problem is given according to the below possible conditions: 1) r_ν is not a multiple eigenvalue or r_ν is a multiple eigenvalue with not more than one linearly independent eigenvector. 2) r_ν is a multiple eigenvlaue with more than one linearly independent eigenvector.

7.2.1 r_ν has only one linearly independent eigenvector

Because $r_\mu = r_\nu$ from (7.1) we shall have

$$0 = \overline{C}^{\nu}_{\nu(j-1)e_\nu} - C^{\nu}_{\nu j e_\nu},$$

$$\overline{C}^{\nu}_{\nu j e_\nu} = \overline{C}^{\nu}_{\nu(j-1)(e_\nu-1)} - C^{\nu}_{\nu j(e_\nu-1)},$$

$$\overline{C}^{\nu}_{\nu j(e_\nu-1)} = \overline{C}^{\nu}_{\nu(j-1)(e_\nu-2)} - C^{\nu}_{\nu j(e_\nu-2)},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\overline{C}^{\nu}_{\nu j(j+1)} = \overline{C}^{\nu}_{\nu(j-1)j} - C^{\nu}_{\nu jj} + \frac{dr_\nu}{d\omega},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\overline{C}^{\nu}_{\nu j2} = \overline{C}^{\nu}_{\nu(j-1)1} - C^{\nu}_{\nu j1}. \tag{7.4}$$

These will yield $\overline{C}^{\nu}_{\nu jk}$ $(1 < k \leq e_\nu)$ except for $\overline{C}^{\nu}_{\nu j1}$, because above equations are in the form

$$\overline{C}^{\nu}_{\nu jk} = \overline{C}^{\nu}_{\nu(j-1)(k-1)} - C^{\nu}_{\nu j(k-1)}, \tag{7.5}$$

except when $k = j+1$ in which case we have

$$\overline{C}^{\nu}_{\nu j(j+1)} = \overline{C}^{\nu}_{\nu(j-1)j} - C^{\nu}_{\nu jj} + \frac{dr_\nu}{d\omega}. \tag{7.6}$$

To find $\overline{C}^{\nu}_{\nu j1}$, the only coefficient left undetermined for the present condition, suppose the arbitrary entries of $f^{(\mu)}_k$ at row g are shown by $f^{(\mu)}_{kg}$. Then for this row g:

$$\frac{df^{(\nu)}_{jg}}{d\omega} = \sum_{k,\mu} \overline{C}^{\mu}_{\nu jk} f^{(\mu)}_{kg} = \left[\sum_{k,\mu} \overline{C}^{\mu}_{\nu jk} f^{(\mu)}_{kg}\right] + \overline{C}^{\nu}_{\nu j1} f^{(\nu)}_{1g}, \tag{7.7}$$

hence

$$\overline{C}^{\nu}_{\nu j1} = \frac{1}{f^{(\nu)}_{1g}} \left[\frac{df^{(\nu)}_{1g}}{d\omega} - \sum_{k,\mu} \overline{C}^{\mu}_{\nu jk} f^{(\mu)}_{kg}\right]. \tag{7.8}$$

where for the summation within brackets in (7.7) and the summation in (7.8) $k = 1$ and $\mu = \nu$ not at the same time.

7.2.2 r_ν has more than one linearly independent eigenvector

Suppose that in (7.1) μ and ν correspond to the generating eigenvectors $f_1^{(s_t)}$ and $f_1^{(s_p)}$ of the same eigenvalue r_s, because $r_\mu = r_\nu = r_s$. From (7.1) we have

$$0 = \overline{C}^{s_t}_{(s_p)(j-1)e_{s_t}} - C^{s_t}_{(s_p)je_{s_t}},$$

$$\overline{C}^{s_t}_{(s_p)je_{s_t}} = \overline{C}^{s_t}_{(s_p)(j-1)(e_{st}-1)} - C^{s_t}_{(s_p)j(e_{st}-1)},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\overline{C}^{s_t}_{(s_p)j(j+1)} = \overline{C}^{s_t}_{(s_p)(j-1)j} - C^{s_t}_{(s_p)jj},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\overline{C}^{s_t}_{(s_p)j2} = \overline{C}^{s_t}_{(s_p)(j-1)1} - C^{s_t}_{(s_p)j1}. \tag{7.9}$$

$\overline{C}^{s_t}_{(s_p)jk}$ can be determined from above relations (7.9) for all k except for $k = 1$, because these equations are in the form

$$\overline{C}^{s_t}_{(s_p)jk} = \overline{C}^{s_t}_{(s_p)(j-1)(k-1)} - C^{s_t}_{(s_p)j(k-1)}. \tag{7.10}$$

Furthermore when in (7.1) μ and ν correspond to the same generating eigenvector $f_1^{(s_i)}$, (7.5) and (7.6) will be replaced by (7.11) and (7.12) below:

$$\overline{C}^{s_i}_{(s_i)jk} = \overline{C}^{s_i}_{(s_i)(j-1)(k-1)} - C^{s_i}_{(s_i)j(k-1)}, \tag{7.11}$$

except when $k = j + 1$ in which case we have

$$\overline{C}^{s_i}_{(s_i)j(j+1)} = \overline{C}^{s_i}_{(s_i)(j-1)j} - C^{s_i}_{(s_i)jj} + \frac{dr_\nu}{d\omega}. \tag{7.12}$$

In (7.11) again $\overline{C}^{s_i}_{(s_i)jk}$ cannot be determined when $k = 1$. To find $\overline{C}^{s_t}_{(s_p)j1}$ consider the expansion $\frac{df_j^{(s_p)}}{d\omega} = \sum_{k,\mu} \overline{C}^{\mu}_{(s_p)jk} f_k^{(\mu)}$ and particularly arbitrary entry rows g of each vector $f_k^{(\mu)}$ which are denoted by $f_{kg}^{(\mu)}$:

$$\frac{df_{jg}^{(s_p)}}{d\omega} = \left[\sum_{k,\mu} \overline{C}^{\mu}_{(s_p)jk} f_{kg}^{(\mu)}\right] + \overline{C}^{s_1}_{(s_p)j1} f_{1g}^{(s_1)} + \overline{C}^{s_2}_{(s_p)j1} f_{1g}^{(s_2)} + \cdots + \overline{C}^{s_{\beta_s}}_{(s_p)j1} f_{1g}^{(s_{\beta_s})}. \tag{7.13}$$

The unknowns $\overline{C}^{s_t}_{(s_p)j1}$ and $\overline{C}^{s_i}_{(s_i)j1}$ are shown outside the summation sign in (7.13). β_s is the geometric multiplicity of r_s. In the summation sign in (7.13) never $\mu = s_t (1 \le t \le \beta_s)$ and $k = 1$ at the same time. When g runs through all possible values we shall have a system of equations to replace (7.13) for the unknowns $\overline{C}^{s_1}_{(s_p)j1}, \overline{C}^{s_2}_{(s_p)j1}, \overline{C}^{s_3}_{(s_p)j1}, \ldots, \overline{C}^{s_{\beta_s}}_{(s_p)j1}$ whose coefficient matrix is F_2^s is as defined in Notation 2.4 of Section 2.

8 Conclusion

Summarizing, for the classical linear eigenvalue problem when the matrix of the problem is dependent on a single real parameter we have found expressions for the derivatives of simple and multiple eigenvalues. We computed the derivatives of eigenvectors of multiple eigenvalues as well as simple eigenvalues for the same problem. In the case of derivatives of generalized eigenvectors we have devised recursion formulae for the expansion coefficients of these derivatives. The formulae relate coefficients of derivatives of consecutively indexed generalized eigenvectors. For the coefficients $\overline{C}^{\mu}_{\nu jk}$ of Eq.(7.1), in Section 7.1 we have determined $\overline{C}^{\mu}_{\nu jk}$ in terms of $\overline{C}^{\mu}_{\nu(j-1)k}$ when $\mu \neq \nu$. Here subscript j stands for the index of a generalized eigenvector. In Section 7.2.1 we have found $\overline{C}^{\nu}_{\nu jk}$ in terms of $\overline{C}^{\nu}_{\nu(j-1)k}$ and in Section 7.2.2 we have expressed $\overline{C}^{s_t}_{(s_p)jk}$ in terms of $\overline{C}^{s_t}_{(s_p)(j-1)k}$ when $f_k^{(s_p)}$ and $f_k^{(s_t)}$ are generalized eigenvectors corresponding repectively to linearly independent generating eigenvectors $f_1^{(s_p)}$ and $f_1^{(s_t)}$ of the same eigenvalue r_s.

References

[1] A. L. Andrew, K.-W. Erich Chu, P. Lancaster, Derivatives of eigenvalues and eigenvectors of matrix functions, *SIAM J. Matrix Anal. Appl.*, **14** (1993), 903-926. https://doi.org/10.1137/0614061

[2] C. Cardani, P. Mantegazza, Calculation of eigenvalue and eigenvector derivatives for algebraic flutter and divergence eigenproblems, *AIAA Journal*, **7** (1979), 408-412. https://doi.org/10.2514/3.61140

[3] R. L. Fox, M. P. Kapoor, Rate of change of eigenvalues and eigenvectors, *AIAA Journal*, **6** (1968), 2426-2429. https://doi.org/10.2514/3.5008

[4] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, Heidelberg, 1966. https://doi.org/10.1007/978-3-642-66282-9

[5] P. Lancaster, On eigenvalues of matrices dependent on a parameter, *Numerische Mathematik*, **6** (1964), 377–387.

[6] L. Padulo, M. A. Arbib, *System Theory*, W. B. Saunders Company, Philadelphia, London, Toronto, 1974.

[7] U. Prells, M.I. Friswell, Partial derivatives of repeated eigenvalues and their eigenvectors, *AIAA Journal*, **8** (1997), 1363-1368.

[8] I. S. Sokolnikoff, R. M. Redheffer, *Mathematics of Physics and Modern Engineering*, McGraw-Hill Book Company, New York, 1966.

[9] J. G. Sun, Multiple eigenvalue sensitivity analysis, *Linear Algebra Appl.*, **137/138** (1990), 183-211.

Received: April 1, 2023, Published: April 18, 2023