Eigenvector Derivatives with Repeated Eigenvalues

R. Lane Dailey*
TRW, Inc., Redondo Beach, California

In this paper an algorithm is derived for computing the derivatives of eigenvalues and eigenvectors for real symmetric matrices in the case of repeated eigenvalues, where the matrices are functions of real parameters such as mass density or moment of inertia. The algorithm is an extension of recent work by I. U. Ojalvo; the key step in this extended derivation is to differentiate the eigenvalue equation twice. The algorithm preserves the symmetry and band structure of the matrices, allowing efficient computer storage and solution techniques. Applications include sensitivity analysis and optimization of the normal modes of finite-element modeled structures, such as large space structures. A cantilever beam finite-element example is included.

Introduction

METHODS for computing the derivatives of matrix eigenvalues and eigenvectors have been studied by many researchers in the past 20 years. The problem is important because these eigensolutions characterize the normal modes of vibration for structures modeled by finite-element methods. Knowing the derivatives of these modes with respect to physical parameters can help an engineer optimize a structure's design or minimize its sensitivity to the parameters.

For structural control systems, these eigenderivatives have direct application for system identification and robust performance tests. With knowledge of the derivatives, an engineer can construct a parameterized evaluation model containing the structure's natural modes. Experimental data can then be used to identify best-fit values for the parameters.[1] Alternatively, one can test whether the closed-loop control system will perform satisfactorily for all parameter values in a given set.[2,3] For structural design, these derivatives can be used to optimize the mode frequencies and mode shapes of a structure by varying its design parameters.

Jacobi[4] first studied the eigenvalue derivative problem in 1846. In the present era, Lancaster[5] treated eigenvalue derivatives in the repeated eigenvalue case. Fox and Kapoor[6] studied eigenvector derivatives for symmetric matrices. Rogers,[7] followed by Plaut and Husseyin,[8] Rudisill,[9] Rudisill and Chu,[10] and Doughty[11] treated the eigenvector derivative problem for nonsymmetric matrices. For a thorough review of the research in sensitivity methods for finite-dimensional structural problems, see the excellent survey paper by Adelman and Haftka.[12]

In 1976, Nelson[13] presented a powerful algorithm for computing the eigenvalue and eigenvector derivatives of general real matrices with nonrepeated eigenvalues. Nelson's method is a significant advance because it requires knowledge of only those eigenvectors that are to be differentiated. Previous methods required computing all or most of the eigenvectors. For the very high-order matrices encountered in finite-element analysis, this greatly affects computer time. Nelson's method also preserves the band structure and symmetry, if any, of the matrices; this allows the use of efficient storage and solution techniques.

Unfortunately, Nelson's method suffers from singularity problems when repeated eigenvalues exist. In typical structures, there are many repeated or nearly equal eigenvalues, due to structural symmetry. Repeated mode frequencies can occur in structures with repeated elements, such as wheelsets on trains. In other cases, frequencies can cross as parameters change; for instance, the torsion and bending mode frequencies of a beam may cross as the length varies.

In a recent paper, Ojalvo[14] presented a method for computing eigenvalue and eigenvector derivatives in the case of real symmetric matrices with repeated eigenvalues. This method is partially incorrect. The eigenvalue derivatives are correctly derived, but the solution for eigenvector derivatives requires solving for m^2 unknowns, where m is the number of repeated eigenvalues. Ojalvo solves for m of these unknowns (the diagonal elements of an $m \times m$ matrix) but then uses an underdetermined system of equations to find the other $m(m-1)$ (off-diagonal elements). This generally gives incorrect results, except in special cases where this $m \times m$ matrix happens to be diagonal. Both examples in Ref. 14 were of this special type.

This paper gives the correct solution for eigenvector derivatives, using the same approach as Ojalvo, but adding additional constraint equations to allow solving for the missing $m(m-1)$ unknowns. The additional constraints are found by differentiating the eigenvalue equation twice. Standard matrix notation is used throughout, which makes the derivation somewhat simpler. As in Nelson's and Ojalvo's methods, the symmetry and band structure of the matrices are preserved, allowing for efficient solution. A finite-element example of a cantilever beam is included.

The first section of this paper reviews Nelson's method[13] for finding eigenvalue and eigenvector derivatives in the case of nonrepeated eigenvalues. The method is derived here for the symmetric case only, although Nelson treats the general nonsymmetric real case. The next section begins by describing Ojalvo's recent method[14] for finding real symmetric matrix eigenvalue and eigenvector derivatives in the repeated eigenvalue case. The deficiency in the method is pointed out, and the new result that corrects the error is then derived. Then the new method is demonstrated for a two-element model of a cantilever beam.

Nelson's Method

First, two definitions: the kernel or nullspace of a matrix K is that set of vectors x such that $Kx = 0$. The range of a matrix K is that set of vectors y such that $y = Kx$ for some vector x.

Given symmetric real matrices K, M, K', and M' in $R^{n \times n}$ where $K' = \partial K/\partial p$, $M' = \partial M/\partial p$ for some real parameter p, let λ in R, x in R^n solve the generalized eigenvalue problem

$$Kx = \lambda Mx \tag{1}$$

Received July 9, 1987; revision received Feb. 1, 1988.

*Member, Technical Staff.

Further assume that λ is a distinct eigenvalue. Also, assume that x is normalized to satisfy

$$x^T M x = 1 \tag{2}$$

In a finite-element structural problem, K is the stiffness matrix and M is the mass matrix. Problem: what are $\lambda' \equiv \partial\lambda/\partial p$ and $x' \equiv \partial x/\partial p$? Differentiate Eqs. (1) and (2):

$$K'x + Kx' = \lambda' Mx + \lambda M'x + \lambda Mx' \tag{3}$$

$$(x')^T Mx + x^T M'x + x^T Mx' = 0 \tag{4}$$

Rearranging Eq. (3) yields

$$(K - \lambda M)x' = (\lambda' M + \lambda M' - K')x \tag{5}$$

Note that transposing Eq. (1) and moving everything to one side of the equation yields $x^T(K - \lambda M) = 0$. Exploiting this fact and the relation $x^T M x = 1$, left-multiply Eq. (5) by x^T:

$$x^T(K - \lambda M)x' = 0 = \lambda' x^T Mx + x^T(\lambda M' - K')x \tag{6}$$

$$\lambda' = x^T(K' - \lambda M')x \tag{7}$$

Now λ' can be used to solve for x'. Define $f \equiv (\lambda' M + \lambda M' - K')x$; then $(K - \lambda M)x' = f$. Note that $(K - \lambda M)$ has rank $n - 1$ and kernel x. This means that for any v solving $(K - \lambda M)v = f$, $v + cx$ is also a solution, where c is any real number. The approach is first to find some particular solution v, then find c so that $v + cx = x'$. A suitable v can be found by setting one element of v to 0 and solving for the remaining elements. This will work for any f in the range of $K - \lambda M$ as long as the corresponding element of x is nonzero. This suggests the following algorithm:

1) Let $G \equiv K - \lambda M$. Assume $x = (x_1, x_2, \ldots, x_n)^T$.
2) Find k such that $|x_k| = \|x\|_\infty \equiv \max_i |x_i|$.
3) Replace the kth row and column of G with zeros except for 1 on the kth diagonal element. Call the result $\bar{G}$.
4) Replace the kth element of f with zero. Call the result $\bar{f}$.
5) Solve $\bar{G}v = \bar{f}$.

Since f is in the range of G (i.e., $x^T f = 0$), the resulting v automatically satisfies $Gv = f$. This process preserves the sparseness and band structure, if any, of G, which allows for efficient solution of $\bar{G}v = \bar{f}$. It remains to find the scalar c. Substitute $x' = v + cx$ into Eq. (4):

$$(cx^T + v^T)Mx + x^T M'x + x^T M(v + cx) = 0 \tag{8}$$

$$c + v^T Mx + x^T M'x + x^T Mv + c = 0 \tag{9}$$

$$c = -v^T Mx - 0.5 x^T M'x \tag{10}$$

This completes the solution for λ' and x'. The complete algorithm is summarized below:

1) Compute $\lambda' = x^T(K' - \lambda M')x$.
2) Let $f \equiv (\lambda' M + \lambda M' - K')x \equiv (f_1, f_2, \ldots, f_n)^T$, and let $G \equiv K - \lambda M$.
3) Find k such that $|x_k| = \|x\|_\infty$ where $x = (x_1, x_2, \ldots, x_n)^T$.
4) Construct $\bar{G}$ by zeroing out row k and column k of G and setting the kth diagonal element to 1.
5) Construct $\bar{f}$ by zeroing out the kth element of f.
6) Solve $\bar{G}v = \bar{f}$.
7) Compute $c = -v^T Mx - 0.5 x^T M'x$.
8) Let $x' = v + cx$.

This is Nelson's method.[13] Note that it has the desirable properties of preserving the structure of M and K (allowing more efficient solution) and of requiring knowledge of only one eigenvalue-eigenvector pair. Both properties are important in realistic structural problems where M and K have very high order.

Ojalvo's Method

Since Nelson's method works only for nonrepeated eigenvalues, we need Ojalvo's method to handle the commonly occurring case of repeated eigenvalues. Suppose we have symmetric K, M, K', and M' as before. Let X in $R^{n \times m}$ be a matrix of eigenvectors solving

$$KX = MX\Lambda \tag{11}$$

where $\Lambda = \lambda I$, and λ in R is the eigenvalue for the eigenspace spanned by the columns of X. That is, λ is an eigenvalue of multiplicity m. Further assume that the columns of X are mass-orthonormal in the sense that

$$X^T M X = I \tag{12}$$

These are the two defining equations in the degenerate (repeated eigenvalues) case. As the parameter p varies, the eigenvalues will, in general, separate. Therefore, the eigenvalue derivative is an m-vector, not a scalar. We will denote it by $\Lambda' = \partial\Lambda/\partial p = \mathrm{diag}(\lambda'_1, \ldots, \lambda'_m)$, a diagonal matrix, since this form simplifies the derivation.

Since the eigenspace is degenerate, any linear combination of the columns of X is an eigenvector. However, the eigenspace splits into as many as m distinct eigenvectors when p varies. For the eigenvector derivative to exist, we must restrict the eigenvectors to lie "adjacent" to the m distinct eigenvectors that appear when p varies. Otherwise, the eigenvectors would jump discontinuously with varying p.

These adjacent eigenvectors can be expressed in terms of X by an orthogonal transformation: $Z = X\Gamma$, where Γ in $R^{m \times m}$ is orthogonal $\Gamma^T\Gamma = I$. The columns of Z are the eigenvectors for which a derivative can be defined. Note that

$$Z^T M Z = \Gamma^T X^T M X \Gamma = \Gamma^T \Gamma = I \tag{13}$$

and so the orthonormality constraint is still satisfied. The problem is to find Γ so that the eigenvector derivatives exist and then to find Z, Z', and Λ'. Differentiate the eigenvalue equation $KZ = MZ\Lambda$:

$$K'Z + KZ' = M'Z\Lambda + MZ'\Lambda + MZ\Lambda' \tag{14}$$

$$(K - \lambda M)Z' = (\lambda M' - K')Z + MZ\Lambda' \tag{15}$$

Premultiply by X^T and substitute $Z = X\Gamma$, noting that $X^T(K - \lambda M) = 0$:

$$X^T(K - \lambda M)Z' = 0 = -X^T(K' - \lambda M')X\Gamma + X^T M X \Gamma \Lambda' \tag{16}$$

$$[X^T(K' - \lambda M')X]\Gamma \equiv D\Gamma = \Gamma\Lambda' \tag{17}$$

This is simply an eigenvalue equation for the matrix $D \equiv X^T(K' - \lambda M')X$. Since $D = D^T$, the columns of Γ will be orthogonal as required. Solving Eq. (17), therefore, yields Γ, $Z = X\Gamma$, and Λ' immediately. It remains to find Z'. Return to Eq. (15), defining $F = (\lambda M' - K')Z + MZ\Lambda'$:

$$(K - \lambda M)Z' \equiv GZ' = F \equiv (\lambda M' - K')Z + MZ\Lambda' \tag{18}$$

Since $K - \lambda M$ has rank $n - m$ and a kernel spanned by the columns of Z, if V is any solution to $GV = F$ then $V + ZC$ is also a solution, where C in $R^{m \times m}$ is an arbitrary matrix. Analogously to Nelson's method, first find some particular solution V, then find C such that $V + ZC = Z'$. This can be done by setting m rows of V to 0 and solving for the remaining elements. This will work for any F in the range of G as long as the corresponding rows of Z (or X) are nonzero. This suggests the following algorithm similar to that used in Nelson's method:

1) Let $Z = [z_{ij}]$.

2) Find k such that z_{k1} is the largest element on the first column of Z.

3) Replace the kth row and column of G with zeros, except place 1 on the kth diagonal element.

4) Replace the kth row of F with zeros.

5) Go back to step 2 and repeat for the next column of Z until through. If k has been used before, choose the second largest (third largest, etc.) element in the column instead.

6) Call the resulting matrices $\bar{G}$ and $\bar{F}$.

7) Solve $\bar{G}V = \bar{F}$.

Since F is known to be in the range of G (i.e., $Z^TF = 0$), the resulting V is guaranteed to solve $GV = F$. It remains to find C so that $V + ZC = Z'$. Differentiate the constraint Eq. (13):

$$(Z')^TMZ + Z^TM'Z + Z^TMZ' = (C^TZ^T + V^T)MZ + Z^TM'Z + Z^TM(V + ZC) = 0 \tag{19}$$

$$C^T + V^TMZ + Z^TM'Z + Z^TMV + C = 0 \tag{20}$$

$$C + C^T = -V^TMZ - Z^TMV - Z^TM'Z \equiv Q \tag{21}$$

Ojalvo suggests letting $C = -0.5Z^TM'Z - Z^TMV$ (using a different notation). This satisfies the constraint on $C + C^T$, but it does not give the correct answer unless C happens to be diagonal. Both examples in Ref. 14 were of this type. The second example in Ref. 14 appears to have nonzero off-diagonal elements in C, but this is due to arithmetic errors; an independent check of the calculations shows that C is diagonal by the above formula in that example.

New Method

The off-diagonal elements of C can be determined by differentiating the eigenvalue equation twice. The original equation is $KZ - MZ\Lambda = 0$; differentiating once gives

$$K'Z + KZ' - M'Z\Lambda - MZ'\Lambda - MZ\Lambda' = 0 \tag{22}$$

Differentiating again yields

$$K''Z + K'Z' + K'Z' + KZ'' - M''Z\Lambda - M'Z'\Lambda - M'Z\Lambda' - M'Z'\Lambda - MZ''\Lambda - MZ'\Lambda' - M'Z\Lambda' - MZ'\Lambda' - MZ\Lambda'' = 0 \tag{23}$$

Combining terms:

$$(K'' - \lambda M'')Z + 2K'Z' + KZ'' - 2M'Z'\Lambda - 2M'Z\Lambda' - MZ''\Lambda - 2MZ'\Lambda' - MZ\Lambda'' = 0 \tag{24}$$

$$(K'' - \lambda M'')Z + 2(K' - \lambda M')Z' + (K - \lambda M)Z'' - 2M'Z\Lambda' - 2MZ'\Lambda' - MZ\Lambda'' = 0 \tag{25}$$

Left-multiply by Z^T and use $Z^T(K - \lambda M) = 0$ and $Z^TMZ = I$:

$$Z^T(K'' - \lambda M'')Z + 2Z^T(K' - \lambda M')Z' - 2Z^TM'Z\Lambda' - 2Z^TMZ'\Lambda' - \Lambda'' = 0 \tag{26}$$

Now expand with $Z' = V + ZC$:

$$Z^T(K'' - \lambda M'')Z + 2Z^T(K' - \lambda M')V + 2Z^T(K' - \lambda M')ZC - 2Z^TM'Z\Lambda' - 2Z^TMV\Lambda' - 2Z^TMZC\Lambda' - \Lambda'' = 0 \tag{27}$$

Now use the relations $Z^T(K' - \lambda M')Z = Z^T[MZ\Lambda' - (K - \lambda M)Z'] = \Lambda'$ and $Z^TMZ = I$:

$$Z^T(K'' - \lambda M'')Z + 2Z^T(K' - \lambda M')V + 2\Lambda'C - 2Z^TM'Z\Lambda' - 2Z^TMV\Lambda' - 2C\Lambda' - \Lambda'' = 0 \tag{28}$$

$$Z^T(K'' - \lambda M'')Z + 2(\Lambda'C - C\Lambda') - \Lambda'' + 2Z^T(K' - \lambda M')V - 2Z^TM'Z\Lambda' - 2Z^TMV\Lambda' = 0 \tag{29}$$

$$C\Lambda' - \Lambda'C + 0.5\Lambda'' = Z^T(K' - \lambda M')V - Z^T(M'Z + MV)\Lambda' + 0.5Z^T(K'' - \lambda M'')Z \equiv R \tag{30}$$

Note that Λ'' is diagonal, whereas $C\Lambda' - \Lambda'C$ always has zeros on the diagonal. This provides a neat separation of C and Λ'' and allows solving for both from this one equation, although only C is of interest.

If $C = [c_{ij}]$, $R = [r_{ij}]$, $\Lambda' = \mathrm{diag}(\lambda'_1,...,\lambda'_m)$ and $\Lambda'' = \mathrm{diag}(\lambda''_1,...,\lambda''_m)$, it is easily seen that

$$r_{ij} = \begin{cases} c_{ij}(\lambda'_j - \lambda'_i) & \text{if } j \neq i \\ 0.5\lambda''_i & \text{if } j = i \end{cases} \tag{31}$$

Therefore, $c_{ij} = r_{ij}/(\lambda'_j - \lambda'_i)$ if $\lambda'_j \neq \lambda'_i$. For $i = j$, c_{ij} is already given by $C + C^T = Q$, and so $c_{ii} = 0.5q_{ii}$. One further problem surfaces when $\lambda'_j = \lambda'_i$ for $i \neq j$, i.e., when there are repeated eigenvalue derivatives. When this occurs, it means that the original m-dimensional eigenspace is not separating into m distinct one-dimensional eigenspaces as p varies; instead, at least some of its dimensions are "sticking together." This could often occur in practice when the parameter p affects the degenerate modes equally or not at all. In these cases, the eigenvector derivative is not unique; any of a continuum of derivatives will move Z into the perturbed eigenspaces as p varies. This provides a free parameter to set arbitrarily: as long as $c_{ij} + c_{ji} = q_{ij} = q_{ji}$, the resulting Z' will be a valid (nonunique) derivative. The simplest choice is to set $c_{ij} = c_{ji} = 0.5q_{ij}$ whenever $\lambda'_i = \lambda'_j$.

The interpretation of the nonunique eigenvector derivatives in Z' that occur with repeated eigenvalue derivatives is analogous to the interpretation of the nonunique eigenvectors that occur with repeated eigenvalues. In the case of repeated eigenvalues, any linear combination z of the corresponding eigenvectors is also an eigenvector if it satisfies the constraint $z^TMz = 1$. In the case of repeated eigenvalue derivatives (for repeated eigenvalues), any corresponding column of $Z' = V + ZC$ is a valid eigenvector derivative if the nonunique elements of C satisfy the constraint $c_{ij} + c_{ji} = q_{ij} = q_{ji}$.

This completes the solution for Z, Z', and Λ'. The complete algorithm is summarized below:

1) Compute $D = X^T(K' - \lambda M')X$.

2) Solve the eigenvalue problem $D\Gamma = \Gamma\Lambda'$. Λ' is the diagonal matrix of eigenvalue derivatives, and Γ should be normalized so that $\Gamma^T\Gamma = I$.

3) Let the columns of $Z = X\Gamma$ be the new eigenvectors.

4) Compute $G = K - \lambda M$, $F = (\lambda M' - K')Z + MZ\Lambda'$.

5) Find the m rows of Z (or X) containing the largest elements. Zero out these rows and columns of G and the same rows of F. Place 1 on the affected diagonal elements of G and call the resulting matrices $\bar{G}$ and $\bar{F}$.

6) Solve $\bar{G}V = \bar{F}$.

7) Compute $Q = C + C^T = -V^TMZ - Z^TMV - Z^TM'Z$.

8) Compute

$$R = C\Lambda' - \Lambda'C + 0.5\Lambda'' = Z^T(K' - \lambda M')V - Z^T(M'Z + MV)\Lambda' + 0.5Z^T(K'' - \lambda M'')Z$$

9) Construct the $m \times m$ matrix C by the rule

$$c_{ij} = \begin{cases} r_{ij}/(\lambda_j' - \lambda_i') & \text{if } \lambda_j' \neq \lambda_i' \\ 0.5 q_{ij} & \text{otherwise} \end{cases} \tag{32}$$

where $\Lambda' = \text{diag}(\lambda_1',\ldots,\lambda_m')$.

10) Let $Z' = V + ZC$. The columns of Z' are the eigenvector derivatives.

Observe that Z' depends on the second derivatives K'' and M'', since these terms are present in the formula for R. It may seem surprising that these second derivatives influence the first derivative Z'. However, it is equally true that the first derivatives K' and M' affect the eigenvectors Z. In this case of repeated eigenvalues, the additional information needed to find Z and Z' comes from the "next level" of derivatives of K and M.

For a great many practical problems, K'' and M'' are zero. For instance, in the cantilever beam example that follows, K'' and M'' would be zero if the design variable (the real parameter p) were chosen to be cross-sectional area, mass density, area moment, or Young's modulus. Choosing beam length as the design parameter, however, would yield nonzero K'' and M''.

Example

This example develops a finite-element model for a cantilever beam, using two finite elements of the type shown in Fig. 1. The elements are combined as in Fig. 2. Because of symmetry, the beam's first mode is a repeated one, involving bending in the y-z plane. Choosing the design parameter p to be the z-axis area moment in the upper element causes only the z-axis component of this double mode to vary. This is reflected in the way the rotation matrix Γ separates the y-axis and z-axis components and in the fact that the y-axis component's derivatives in Λ' and Z' are zero.

For the 8-degree-of-freedom beam element in Fig. 1, the stiffness and mass matrices K_e and M_e are given[15] by

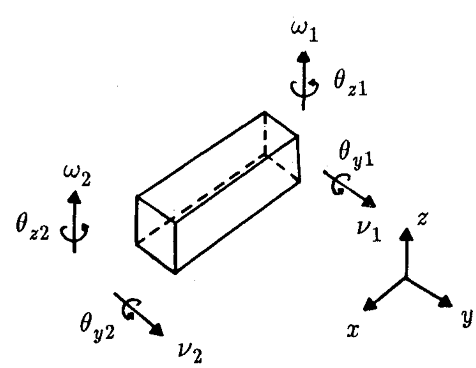

Fig. 1 Beam finite element.

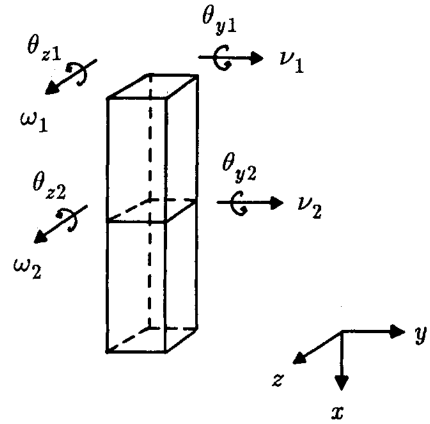

Fig. 2 Two-element cantilever beam.

$$K_e = \frac{E}{L}\begin{bmatrix} 12I_z/L^2 & 0 & 0 & 6I_z/L & -12I_z/L^2 & 0 & 0 & 6I_z/L \\ & 12I_y/L^2 & -6I_y/L & 0 & 0 & -12I_y/L^2 & -6I_y/L & 0 \\ & & 4I_y & 0 & 0 & 6I_y/L & 2I_y & 0 \\ & & & 4I_z & -6I_z/L & 0 & 0 & 2I_z \\ & \text{symmetric} & & & 12I_z/L^2 & 0 & 0 & -6I_z/L \\ & & & & & 12I_y/L^2 & 6I_y/L & 0 \\ & & & & & & 4I_y & 0 \\ & & & & & & & 4I_z \end{bmatrix} \tag{33}$$

$$M_e = \frac{\rho AL}{420}\begin{bmatrix} 156 & 0 & 0 & 22L & 54 & 0 & 0 & -13L \\ & 156 & -22L & 0 & 0 & 54 & 13L & 0 \\ & & 4L^2 & 0 & 0 & -13L & -3L^2 & 0 \\ & & & 4L^2 & 13L & 0 & 0 & -3L^2 \\ & \text{symmetric} & & & 156 & 0 & 0 & -22L \\ & & & & & 156 & 22L & 0 \\ & & & & & & 4L^2 & 0 \\ & & & & & & & 4L^2 \end{bmatrix} \tag{34}$$

where E is Young's modulus, L is the beam length, A is cross-sectional area, ρ is mass density, and I_y and I_z are area moments for the y and z axes. The discrete displacement vector is

$$q_e = [v_1 \ w_1 \ \theta_{y1} \ \theta_{z1} \ v_2 \ w_2 \ \theta_{y2} \ \theta_{z2}]^T \tag{35}$$

Now combine two elements as in Fig. 2, with displacement and bending constrained at the base to form a cantilever beam. The global displacement vector is now

$$q = [v_1 \ w_1 \ \theta_{y1} \ \theta_{z1} \ v_2 \ w_2 \ \theta_{y2} \ \theta_{z2}]^T \tag{36}$$

The global stiffness and mass matrices are

$$K_g = B_1^T K_{e1} B_1 + B_2^T K_{e2} B_2 \tag{37}$$

$$M_g = B_1^T M_{e1} B_1 + B_2^T M_{e2} B_2 \tag{38}$$

where

$$B_1 = I \qquad B_2 = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \tag{39}$$

are 8×8 matrices relating the element displacement vectors $\boldsymbol{q}_{e1}$, $\boldsymbol{q}_{e2}$ to the global vector $\boldsymbol{q}$: $\boldsymbol{q}_{e1} = B_1\boldsymbol{q}$, $\boldsymbol{q}_{e2} = B_2\boldsymbol{q}$. Each block of B_2 is a 4×4 submatrix. Given the derivatives K'_{e1}, K'_{e2}, M'_{e1}, and M'_{e2}, Eqs. (37) and (38) can be used to find the global matrix derivatives K'_g and M'_g. Differentiate K_e and M_e with respect to the z-axis area moment I_z:

$$K'_e = \frac{\partial K_e}{\partial I_z} = \frac{E}{L}\begin{bmatrix} 12/L^2 & 0 & 0 & 6/L & -12/L^2 & 0 & 0 & 6/L \\ & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 4 & -6/L & 0 & 0 & 2 \\ & \text{symmetric} & & & 12/L^2 & 0 & 0 & -6/L \\ & & & & & 0 & 0 & 0 \\ & & & & & & 0 & 0 \\ & & & & & & & 4 \end{bmatrix} \tag{40}$$

$$M'_e = \frac{\partial M_e}{\partial I_z} = 0 \tag{41}$$

Now let I_z vary in the upper element only, so that K'_{e1} is given by Eq. (40) and $K'_{e2} = 0$. Also, set $L = I_z = I_y = \rho = 1$, $A = 420$, and $E = 1000$. The resulting values for K_g, M_g, K'_g, and M'_g are

$$K_g = \begin{bmatrix} 12 & 0 & 0 & 6 & -12 & 0 & 0 & 6 \\ & 12 & -6 & 0 & 0 & -12 & -6 & 0 \\ & & 4 & 0 & 0 & 6 & 2 & 0 \\ & & & 4 & -6 & 0 & 0 & 2 \\ & \text{symmetric} & & & 24 & 0 & 0 & 0 \\ & & & & & 24 & 0 & 0 \\ & & & & & & 8 & 0 \\ & & & & & & & 8 \end{bmatrix} \times 1000 \tag{42}$$

$$M_g = \begin{bmatrix} 156 & 0 & 0 & 22 & 54 & 0 & 0 & -13 \\ & 156 & -22 & 0 & 0 & 54 & 13 & 0 \\ & & 4 & 0 & 0 & -13 & -3 & 0 \\ & & & 4 & 13 & 0 & 0 & -3 \\ & \text{symmetric} & & & 312 & 0 & 0 & 0 \\ & & & & & 312 & 0 & 0 \\ & & & & & & 8 & 0 \\ & & & & & & & 8 \end{bmatrix} \tag{43}$$

$$K'_g = \begin{bmatrix} 12 & 0 & 0 & 6 & -12 & 0 & 0 & 6 \\ & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 4 & -6 & 0 & 0 & 2 \\ & \text{symmetric} & & & 12 & 0 & 0 & -6 \\ & & & & & 0 & 0 & 0 \\ & & & & & & 0 & 0 \\ & & & & & & & 4 \end{bmatrix} \times 1000 \tag{44}$$

and $M'_g = 0$. Furthermore, $K''_g = M''_g = 0$. Solving $K_g X = \lambda M_g X$ yields four equal-frequency pairs of modes because of the structure's y-z symmetry. The lowest frequency pair occurs at $\lambda = 1.8414$, or 0.2160 Hz, with eigenvectors

$$X = \begin{bmatrix} +0.0440 & -0.0533 \\ -0.0533 & -0.0440 \\ -0.0367 & -0.0303 \\ -0.0303 & +0.0367 \\ +0.0149 & -0.0181 \\ -0.0181 & -0.0149 \\ -0.0310 & -0.0256 \\ -0.0256 & +0.0310 \end{bmatrix} \tag{45}$$

Note that the columns of X are not unique, since λ is a repeated eigenvalue. This choice for X satisfies $X^T M_g X = I$. Now apply Ojalvo's modified method: first construct $D = X^T(K'_g - \lambda M'_g)X$ and solve the eigenvalue problem $D\Gamma = \Gamma\Lambda'$. This yields

$$\Lambda' = \begin{bmatrix} 0 & 0 \\ 0 & 0.091856 \end{bmatrix} \qquad \Gamma = \begin{bmatrix} -0.7709 & +0.6369 \\ -0.6369 & -0.7709 \end{bmatrix} \tag{46}$$

The new eigenvector pair is

$$Z = X\Gamma = \begin{bmatrix} 0 \\ +0.0691 \\ +0.0475 \\ 0 \\ 0 \\ +0.0235 \\ +0.0402 \\ 0 \end{bmatrix} \begin{bmatrix} +0.0691 \\ 0 \\ 0 \\ -0.0475 \\ +0.0235 \\ 0 \\ 0 \\ -0.0402 \end{bmatrix} \tag{47}$$

The eigenvector derivative can be defined only for this choice of eigenvectors. Continuing, construct $G = K_g - \lambda M_g$ and $F = (\lambda M'_g - K'_g)Z + M_g Z \Lambda'$. Since the largest components of Z occur in rows 1 and 2, zero out the first and second rows and columns of G and the first and second rows of F, while placing 1 on the first and second diagonal elements of G. Call the results $\bar{G}$ and $\bar{F}$ and solve $\bar{G}V = \bar{F}$.

It remains to find C so that $Z' = V + ZC$. Recall that $K''_g = M''_g = 0$. Compute

$$Q = C + C^T = -V^T M_g Z - Z^T M_g V - Z^T M'_g Z \tag{48}$$

$$R = C\Lambda' - \Lambda' C + 0.5\Lambda'' = Z^T(K'_g - \lambda M'_g)V - Z^T(M'_g Z + M_g V)\Lambda' \tag{49}$$

Constructing C from Q and R by the formula in Eq. (32) and computing Z' yields

$$Z' = V + ZC = \begin{bmatrix} 0 & -0.002052 \\ 0 & 0 \\ 0 & 0 \\ 0 & +0.005109 \\ 0 & +0.001200 \\ 0 & 0 \\ 0 & 0 \\ 0 & -0.001999 \end{bmatrix} \tag{50}$$

The accuracy of these derivatives was tested by a perturbation and differencing scheme. Perturbed matrices $K_p = K_g + 0.001K'_g$, $M_p = M_g + 0.001M'_g = M_g$ were constructed and solved for $K_p X_p = M_p X_p \Lambda_p$, with constraint $X_p^T M_p X_p = I$. The derivatives were estimated by $Z' \approx (X_p - Z)/0.001$, $\Lambda' \approx (\Lambda_p - \lambda I)/0.001$. The results differ from the values of Z' and Λ' computed above by only 0.09% for Λ' and by no more than 0.10% in any nonzero element of Z'.

Note that whereas X shows no separation of y-axis and z-axis motion, $Z = X\Gamma$ clearly divides the eigenspace into y-axis and z-axis components. This is intuitively reasonable, since raising the z-axis moment I_z will increase the frequency of the z-axis mode while not affecting the y-axis mode. A derivative cannot be defined for the eigenvectors in X, since an infinitesimal change in I_z would cause the mode shapes to separate as in Z.

Conclusion

This paper presents a method for computing the derivatives of eigenvalues and eigenvectors when there are repeated eigenvalues. The method corrects a deficiency in Ojalvo's method[14] caused by an underdetermined system of equations. The deficiency was removed by differentiating the eigenvalue equation $KZ = MZ\Lambda$ twice, yielding enough additional constraints to solve the system of equations.

Except for requiring the solution of a low-order eigenvalue problem (with order equal to the number of repeated eigenvalues), the method requires only matrix-vector multiplication and addition and the solution of one high-order linear equation. Like Nelson's method, this method preserves the symmetry, band structure, and sparseness of M and K, and so the linear equation solution is efficient.

In addition to the derivatives Z' and Λ', the method finds a matrix Z whose columns are eigenvectors for which derivatives can be defined; these can be useful in themselves. In the example, the original X shows modes with components in both y and z axes; the vectors returned in Z show a clear separation of these axes, since the stiffness in only one axis is being varied.

It is important to realize that if the derivatives Z' and Λ' are computed with respect to more than one parameter, the eigenvectors in Z will not necessarily be the same for all parameters. That is, the differentiable eigenvectors do not match from one design variable to another. This complicates the construction of multivariable Taylor series for structural design optimization. However, it is necessary in order for the eigenvector derivatives to exist.

One possible application for this new capability is in structural control system sensitivity analysis. Given the derivatives of the eigenvalues and eigenvectors of the most important modes of a structure, it should be possible to compute the derivative of the structure's transfer-function matrix. This could be used to analyze the robustness of a controller design and also to generate plant uncertainty models for H_∞ or structured singular value (SSV) controller synthesis.[3] Since the parameters are assumed to be real, the new exact methods for computing multivariable stability margin[2] could be applied. The derivatives could also be used to perform best-fit identification of parameters from experimental data.[1] Another application is structural design optimization. These derivatives could be used to optimize the mode frequencies and mode shapes of a structure by varying its design parameters.

References

[1]Lukich, M. S. and Major, C. S., "Application of Experimentally Identified Parameters to Control of a Large Flexible Structure," AIAA GNC Conference, Snowmass, CO, Aug. 1985.

[2]de Gaston, R. R. E. and Safonov, M. G., "Exact Calculation of the Multiloop Stability Margin," *IEEE Transactions on Automatic Control*, Vol. AC-33, Feb. 1988, pp. 156–171.

[3]Doyle, J. C., Lecture notes, ONR/Honeywell Workshop on Advances in Multivariable Control, Minneapolis, MN, Oct. 1984.

[4]Jacobi, K. G. J., "Uber ein leichtes Verfahren die in der Theorie der Saecularstoerungen vorkommenden Gleichungen numerisch aufzuloesen," *Zeitschrift fur Reine und Angewandte Mathematik*, Vol. 30, 1846, pp. 51–95; NASA TT.F-13,666, June 1971.

[5]Lancaster, P., "On Eigenvalues of Matrices Dependent on a Parameter," *Numerische Mathematik*, Vol. 6, No. 5, 1964, pp. 377–387.

[6]Fox, R. L. and Kapoor, M. P., "Rate of Change of Eigenvalues and Eigenvectors," *AIAA Journal*, Vol. 6, Dec. 1968, pp. 2426–2429.

[7]Rogers, L. C., "Derivatives of Eigenvalues and Eigenvectors," *AIAA Journal*, Vol. 8, May 1970, pp. 943–944.

[8]Plaut, R. H. and Husseyin, K., "Derivatives of Eigenvalues and Eigenvectors in Non-Self-Adjoint Systems," *AIAA Journal*, Vol. 11, Feb. 1973, pp. 250–251.

[9]Rudisill, C. S., "Derivatives of Eigenvalues and Eigenvectors of a General Matrix," *AIAA Journal*, Vol. 12, May 1974, pp. 721–722.

[10]Rudisill, C. S. and Chu, Y., "Numerical Methods for Evaluating the Derivatives of Eigenvalues and Eigenvectors," *AIAA Journal*, Vol. 13, June 1975, pp. 834–837.

[11]Doughty, S., "Eigenvalue Derivatives of Damped Torsional Vibrations," *ASME Journal of Mechanical Design*, Vol. 104, April 1982, pp. 463–465.

[12]Adelman, H. M. and Haftka, R. T., "Sensitivity Analysis of Discrete Structural Systems," *AIAA Journal*, Vol. 24, May 1986, pp. 823–832.

[13]Nelson, R. B., "Simplified Calculations of Eigenvector Derivatives," *AIAA Journal*, Vol. 14, Sept. 1976, pp. 1201–1205.

[14]Ojalvo, I. U. "Gradients for Large Structural Models with Repeated Frequencies," Aerospace Technology Conference and Exposition, Society of Automotive Engineers, Warrendale, PA, Paper 86-1789.

[15]Lukich, M. S., "Structural Dynamic Model Optimization for Control System Design," Ph.D. Thesis, Dept. of Mechanical, Aerospace, and Nuclear Engineering, UCLA, Los Angeles, CA, 1987.