A Discourse on Sensitivity Analysis for Discretely-Modeled Structures

Howard M. Adelman and Raphael T. Haftka

March 1991
A Discourse on Sensitivity Analysis
for Discretely Modeled Structures

Howard M. Adelman
NASA Langley Research Center
Hampton, Virginia
and
Raphael T. Haftka
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Introduction

Sensitivity analysis is emerging as a fruitful area of engineering research. The reason for this interest is the recognition of the variety of uses for sensitivity derivatives. In its early stages, sensitivity analysis found its predominant use in assessing the effect of varying parameters in mathematical models of control systems; see, for example, the texts of Tomovic, Brayton and Spence, Frank, and Radanovic for discussions of the early development of sensitivity theory. Interest in optimal control in the early 1960s (see, for example, Ref. 5) and automated structural optimization (see for example, Ref. 6) led to the use of gradient-based mathematical programming methods in which derivatives were used to find search directions toward optimum solutions. More recently, there has been strong interest in promoting systematic structural optimization as a useful tool for the practicing structural design engineer on large problems—a process still under way. Early attempts to use formal optimization for large structural systems resulted in excessively long and expensive computer runs. Examination of the optimization procedures indicated that the predominant contributor to the cost and time was the calculation of derivatives. As a consequence, emerging interest in sensitivity analysis has emphasized efficient computational procedures. In addition, researchers have developed and applied sensitivity analysis for approximate analysis, analytical model improvement, and assessment of design trends—so that structural sensitivity analysis has become more than a utility for optimization and is a versatile design tool in its own right. Most recently, researchers in disciplines such as physiology, thermodynamics, physical chemistry, and aerodynamics have been using sensitivity methodology to assess the effects of parameter variations in their analytical models and to create designs insensitive to parameter variation.

Derivatives of structural response can be calculated analytically at three stages. We
can differentiate the continuum equations defining the response of the structure. We can differentiate the equations obtained when the continuum equations are discretized which is the topic of the present chapter. Finally, we can differentiate directly the computer program used to solve the structural response, such as a finite element computer program. This third approach is not discussed in this textbook, but the interested reader is referred to Refs 15 and 16. Analytical derivative calculations typically entail a substantial effort of analysis and software development. In many cases it is better to use derivatives obtained from a finite difference approximation. This chapter therefore starts with the discussion of the calculation of derivatives by finite differences.

**Finite Difference Sensitivities**

**Truncation and Condition Errors**

The simplest finite difference approximation is the first-order forward-difference approximation. Given a function \( g(v) \) of a design variable \( v \), the forward-difference approximation \( \Delta g/\Delta v \) to the derivative \( dg/dv \) is given as

\[
\frac{\Delta g}{\Delta v} = \frac{g(v + \Delta v) - g(v)}{\Delta v}
\]  

(1)

Another commonly used finite-difference approximation is the second-order central-difference approximation

\[
\frac{\Delta g}{\Delta v} = \frac{g(v + \Delta v) - g(v - \Delta v)}{2\Delta v}
\]  

(2)

It is also possible to employ higher-order finite-difference approximations, but they are rarely used in structural optimization applications because of the associated high computational cost. If we need to find the derivatives of the structural response with respect to \( n \) design variables the forward-difference approximation requires \( n \) additional analyses, the central-difference approximation \( 2n \) additional analyses, and higher order approximations are even more expensive.

The key to the selection of the approximation and the step size \( \Delta x \) is an estimate of the required accuracy. This topic is discussed in Ref. 17, and is summarized next.

Whenever finite-difference formulae are used to approximate derivatives, there are two sources of error: truncation and condition errors. The truncation error \( e_T(\Delta v) \) is a result of the neglected terms in the Taylor series expansion of the perturbed function. For
example, the Taylor series expansion of \(g(v + \Delta v)\) can be written as

\[
g(v + \Delta v) = g(v) + \Delta v \frac{\Delta g}{\Delta v}(v) + \frac{(\Delta v)^2}{2} \frac{d^2 g}{dv^2}(v + \zeta \Delta v), \quad 0 \leq \zeta \leq 1 \quad (3)
\]

From Eq. (3) it follows that the truncation error for the forward-difference approximation is

\[
e_T(\Delta v) = \frac{\Delta v}{2} \frac{d^2 g}{dv^2}(v + \zeta \Delta v) \quad 0 \leq \zeta \leq 1 \quad (4)
\]

Similarly, by including one more term in the Taylor series expansion we get that the truncation error for the central difference approximation is

\[
e_T(\Delta v) = \frac{\Delta v^2}{6} \frac{d^3 g}{dv^3}(v + \zeta \Delta v) \quad 0 \leq \zeta \leq 1 \quad (5)
\]

The condition error is the difference between the numerical evaluation of the function and its exact value. One contribution to the condition error is round-off error in calculating \(dg/dv\) from the original and perturbed values of \(g\). This contribution is comparatively small for most computers unless \(\Delta v\) is extremely small. However if \(g(x)\) is computed by a lengthy or ill-conditioned numerical process, the round-off contribution to the condition error can be substantial. Additionally, condition errors may result if \(g(x)\) is calculated by an iterative process which is terminated early. If we have a bound \(\epsilon_g\) on the absolute error in the computed function \(g\), we can estimate the condition error. For example, for the forward-difference approximation the condition error \(e_C(\Delta v)\) is (very!) conservatively estimated from Eq. (1) as

\[
e_C(\Delta v) = \frac{2}{\Delta v} \epsilon_g \quad (6)
\]

Equations (4) and (6) present us with the so called "step-size dilemma." If we select the step size to be small, so as to reduce the truncation error, we may have an excessive condition error. In some cases there may not be any step size which results in acceptable error!

A bound \(e\) on the total error, the sum of the truncation and condition errors, for the forward-difference approximation is obtained from Eqs. (4) and (6) as

\[
e = \frac{\Delta v}{2} |s_b| + \frac{2}{\Delta v} \epsilon_g \quad (7)
\]
where $s_b$ is a bound on the second derivative in the interval $[v, v + \Delta v]$. When $\epsilon_g$ and $s_b$ are available it is possible to calculate an optimum step-size that minimizes $\epsilon$ as

$$\Delta v_{opt} = 2 \sqrt{\frac{\epsilon_g}{|s_b|}}$$

(8)

Procedures for estimating $s_b$ and $\epsilon_u$ are given in Ref 17 and 18.

Iteratively solved Problems

Condition errors can become important when iterative methods are used for performing some of the calculations. Consider a simple example of a single displacement component $u$ which is obtained by solving a nonlinear algebraic equation which depends on one design variable $v$

$$f(v, u) = 0$$

(9)

The solution of Eq. (9) is obtained by an iterative process which starts with some initial guess of $u$ and terminates when the iterant $\tilde{u}$ is estimated to be within some tolerance $\epsilon$ of the exact $u$ (Note that $\epsilon$ is a bound on the condition error in $u$). To calculate the derivative $du/dv$ assume that we use the forward-difference approximation. That is, we perturb $v$ by $\Delta v$ and solve Eq. (9) for $u_\Delta$

$$f(v + \Delta v, u_\Delta) = 0$$

(10)

The iterative solution of Eq.(10) yields an approximation $\tilde{u}_\Delta$, and then $du/dv$ is approximated as

$$\frac{du}{dv} \approx \frac{\tilde{u}_\Delta - \tilde{u}}{\Delta v}$$

(11)

To start the iterative process for obtaining $u_\Delta$, two initial guesses come to mind. The first is to start with the same initial guess that was used to solve for $u$. If the convergence of the iterative process is monotonic there is a good chance that when we use Eq. (11) the errors in $\tilde{u}$ and $\tilde{u}_\Delta$ will almost cancel out, and we will get a very small condition error. The other logical initial guess for $u_\Delta$ is $\tilde{u}$. This initial guess is known to be good because $\Delta x$ is typically small, and so we may get fast convergence. Unfortunately, this time we cannot expect the condition errors to cancel. As we iterate on $\tilde{u}_\Delta$, the original error (the difference between $u$ and $\tilde{u}$) will be reduced at the same time that the change due to $\Delta x$ is taking effect (consider, for example, what happens if $\Delta x$ is set to zero, or an extremely small number).

Reference 19 suggests a strategy which allows us to start the iteration for $u_\Delta$ from $\tilde{u}$ without worry of excessive condition errors. The approach is to pretend that $\tilde{u}$ is the exact
rather than approximate solution by changing the problem that we want to solve. Indeed, \( \bar{u} \) is the exact solution of
\[
f(v, u) - f(v, \bar{u}) = 0	ag{12}
\]
which is only slightly different from our original problem (because \( f(v, \bar{u}) \) is almost zero). We now find the derivative \( du/dv \) from Eq. (12), by obtaining \( u_\Delta \) as the solution of
\[
f(v + \Delta v, u_\Delta) - f(v, \bar{u}) = 0
\]
Because \( \bar{u} \) is the exact solution of this equation for \( \Delta v = 0 \) the iterative process will only reflect the effect of \( \Delta v \), and we will obtain a good approximation from Eq. (11).

Because of the high cost and the accuracy problems associated with finite-difference derivatives, there has been much effort into developing analytical derivative approximations. The rest of this chapter is devoted to such analytical expressions for sensitivity derivatives.

**Sensitivity of Static Response**

**First Derivatives of Linear Response**

This section of the paper focuses on the calculation of first derivatives of static linear structural response (displacements and stresses) computed from finite element models. The governing equation for displacement is
\[
K U = F
\]
where \( K \) is the symmetric stiffness matrix of order \( n \times n \), \( U \) the vector of displacements, and \( F \) the vector of applied forces. Both \( K \) and \( F \) are, in general, functions of design variables \( v \). A typical function of displacement (e.g., a constraint) will be represented as
\[
g = g(U, v), \quad U = U(v)	ag{15}
\]
Analytical calculations of derivatives of displacements and their functions have been performed by three methods: the direct or design space method, the atjoint variable or state space method, and the virtual load method. The virtual load method is a special case of the direct method. Both the direct and adjoint methods begin with the differentiation of Eqs. (14) and (15).
\[
K \frac{dU}{dv} = \frac{\partial F}{\partial v} - \frac{\partial K}{\partial v} U \equiv R_v
\]

The direct method is to solve Eq. (16) for \( \frac{dU}{dv} \) and substitute \( \frac{dU}{dv} \) into Eq. (17). Equation (16) needs to be solved once for each design variable \( v \) so that the direct method is costly when the number of design variables is large.

The adjoint variable or state space method starts by defining a vector of adjoint variables that satisfies the equation

\[
K\lambda = \frac{\partial g}{\partial U}
\]

where \( \frac{\partial g}{\partial U} \) is sometimes referred to as the dummy load vector. (If \( g \) is a particular displacement component, then \( \frac{\partial g}{\partial U} \) corresponds to a force of unit magnitude in the direction of the component.) Then using Eqs. (16-18),

\[
\frac{dg}{dv} = \frac{\partial g}{\partial v} + \lambda^T R_v
\]

The adjoint variable method requires the solution of Eq. (18) once for each function \( g \). Therefore, if the number of functions is smaller than the number of design variables, the adjoint variable method is more efficient and, conversely, if the number of design variables is smaller, the direct approach is more efficient. Both the direct and adjoint methods involve fewer computations than the finite difference approach, which requires repeated factorization of the stiffness matrix, whereas the direct and adjoint methods require a single factorization with several right-hand sides.

Calculation of \( \partial K/\partial v \)

An important computational task in the adjoint and direct methods is the calculation of \( \partial K/\partial v \). If the structural model contains only elements whose stiffness matrix is proportional to \( v \) (such as rods where \( v \) is the cross-sectional area or membranes and shear panels where \( v \) is the thickness), \( \partial K/\partial v \) is a constant matrix. But for elements having bending stiffness such as beams and plates, the stiffness matrix is a nonlinear function of the cross-sectional dimensions, and the stiffness matrix derivatives are not easily evaluated. The difficulties associated with shape design variables are even more severe. Analytical expressions for derivatives of the stiffness matrix are cumbersome and more expensive to evaluate than the stiffness matrix itself. Furthermore, coding analytical derivatives of
stiffness matrices with respect to all possible design variables is a formidable task, especially that in many cases users of structural analysis software that does not have sensitivity capabilities do not have access to the source code of the software. For these reasons, the preferred approach by most analysts is to compute $\partial K / \partial \nu$ by finite differences. This combination of analytical derivative expressions such as Eq. (19) coupled with finite-difference evaluation of the stiffness matrix is known as the semi-analytical method.

Unfortunately, the semi-analytical method is prone to large errors for some shape design variables. The problem was explained in Ref. 20 by noting that Eq. (16) treats the sensitivity of the displacement vector as the solution of a structural analysis problem with the load replaced by the pseudo-load vector $R_\nu$. This presupposes that the derivative of the displacement vector is a legitimate displacement vector itself, which is not always the case. A simple example when the derivative of the displacement vector is not a legitimate displacement is a nearly incompressible material (Poisson's ratio close to 0.5). The derivative of the displacement with respect to shape changes, treated as a displacement field, would typically represent large volume changes. Thus the pseudo-load vector, $R_\nu$, would need to have extremely large components to extract such large volume changes from a nearly incompressible material. In such a case, the small truncation errors associated with the finite-difference calculation of the pseudo-load are greatly amplified with a resulting very poor accuracy of the semi-analytical sensitivities. A similar phenomenon can occur for shape changes in bending problems, such as those associated with beams, plates and shells. The sensitivity field is often dominated by shear deformations. Since it is very difficult to force a slender beam or a thin shell to undergo large shearing deformations, we again require very large pseudo loads with disastrous effects of small errors in these loads.

**Calculation of Second Derivatives**

Second derivatives of displacement and constraint functions are used for approximate analysis, and for the calculation of derivatives of optimal solutions. Such derivatives may be obtained by differentiating Eqs. (16) and (17), for example,

$$K \frac{d^2 U}{d\nu^2} = \frac{\partial R_\nu}{\partial \nu} + \frac{\partial R_\nu}{\partial U} \frac{dU}{d\nu}$$

$$\frac{d^2 g}{d\nu^2} = \frac{\partial^2 g}{\partial \nu^2} + 2 \left( \frac{\partial^2 g}{\partial U \partial \nu} \right)^T \frac{dU}{d\nu} + \left( \frac{\partial g}{\partial U} \right)^T \frac{d^2 U}{d\nu^2}$$

(20)
However, for \( m \) design variables there are \( m(m + 1)/2 \) second derivatives, and Eqs. (20) need to be solved for that many right-hand sides. It is possible to proceed with a more efficient approach to use Eq. (18) to obtain

\[
\frac{d^2 g}{dv^2} = \frac{\partial^2 g}{\partial v^2} + \left( \frac{\partial^2 g}{\partial U \partial v} \right) \frac{dU}{dv} + \lambda^T \left( \frac{\partial R_v}{\partial v} + \frac{\partial R_v}{\partial U} \frac{dU}{dv} \right)
\]  

(21)

This approach requires the solution of Eq. (16) for all the first derivatives and Eq. (6) for all vectors of adjoint variables.

**Stress Derivatives**

The stresses in an element may be obtained from the displacements using

\[
\sigma = SU - GT
\]  

(22)

where \( \sigma \) is a vector of element stresses, \( T \) is an element temperature, and \( S \) and \( G \) are stress-displacement and stress-temperature matrices, respectively.

Derivatives of stresses may be obtained by differentiating Eq. (22).

\[
\frac{d\sigma}{dv} = S \frac{dU}{dv} + \frac{\partial S}{\partial v} U - \frac{\partial G}{\partial v} T
\]  

(23)

For finite elements such as rods, membranes, and shear panels, \( S \) and \( G \) are independent of \( v \), and stress derivatives are obtained by simply substituting \( dU/dv \) for \( U \) and \( T = 0 \) in Eq. (22). For bending-type elements, \( S \) and \( G \) may be functions of \( v \) and the complete expression must be used; see Camarda and Adelman.\(^{21}\)

**Derivatives of Nonlinear Response**

In the case of nonlinear analysis, the equations of equilibrium may be written as

\[
P(U, v) = \mu F(v)
\]  

(24)

where \( P \) is the internal force generated by the deformation of the structure, and \( \mu F \) is the external applied load. The load scaling factor \( \mu \) is typically used in nonlinear analysis procedures for tracking the evolution of the solution as the load is increased. This is useful because the equations of equilibrium may have several solutions for the same applied loads.
By increasing $\mu$ gradually we make sure that we obtain the solution that corresponds to the structure being loaded from zero.

Differentiating Eq. (24) with respect to the design variable $v$ we obtain

$$J \frac{dU}{dv} = \mu \frac{dF}{dv} - \frac{\partial F}{\partial v}$$

(25)

where $J$ is the Jacobian of $P$ at $U$,

$$J_{kl} = \frac{\partial F_k}{\partial U_i}$$

(26)

often called the tangential stiffness matrix.

The direct method for obtaining $dg/dv$ is to solve Eq. (25) for $dU/dv$ and substitute into Eq. (17). The matrix $J$ is often available from the solution of the equations of equilibrium when these are solved by using Newton’s method. Newton’s method is based on a linear approximation of the equations of equilibrium about a trial solution $\bar{U}$

$$P(\bar{U}, v) + J(\bar{U}, v)(U - \bar{U}) \approx \mu F(v)$$

(27)

Equation (27) solved for $U$, typically provides a better approximation to $U$ than $\bar{U}$. This new approximation replaces $\bar{U}$ in Eq. (27) for the next iteration, either with an updated value of $J$ (Newton’s method) or with the old value (modified Newton’s method). The iteration continues until convergence to a desired accuracy is achieved. If the last iterate $\bar{U}$, for which $J$ was calculated, is close enough to $U$, then that $J$ can be used for calculating the derivative of $U$.

The adjoint approach is very similar to that used in the linear case. The adjoint vector $\lambda$ is the solution of the equation

$$J^T \lambda = \frac{\partial g}{\partial U}$$

(28)

Then it is easy to check that we obtain

$$\frac{dg}{dv} = \frac{\partial g}{\partial v} + \lambda^T (\mu \frac{\partial F}{\partial v} - \frac{\partial P}{\partial v})$$

(29)

Sensitivity of Eigenvalues and Eigenvectors
Distinct Eigenvalues

The general problem is to compute derivatives of eigenvalues and eigenvectors with respect to design variables or system parameters. For reference purposes, the most general case considered is the following eigenvalue problem:

\[ AX = \lambda BX \]  \hspace{1cm} (30)

\[ Y^T A = \lambda Y^T B \]  \hspace{1cm} (31)

\[ Y^T BX = 1 \]  \hspace{1cm} (32)

where \( \lambda \) is an eigenvalue (generally complex). The generally nonsymmetric real \( n \times n \) matrices \( A \) and \( B \) are assumed to be explicit functions of a set of design variables \( v \), and \( X \) and \( Y \) are right and left eigenvectors, respectively. The first result on eigenvalue derivatives was published by Jacobi,\(^{22}\) who developed the result for the special case of symmetric \( A \), and \( B = I \)

\[ \frac{\partial \lambda}{\partial v} = Y^T \frac{\partial A}{\partial v} X \]  \hspace{1cm} (33)

Wittrick\(^{23}\) applied Jacobi’s formula for the case of a symmetric matrix to the derivatives of buckling eigenvalues and presented results for the change in buckling loads of plates with respect to aspect ratio and thickness.

Fox and Kapoor\(^{24}\) and Fox\(^{25}\) considered the special case of symmetric \( A \) and \( B \) matrices. For eigenvalues their formula is

\[ \frac{\partial \lambda}{\partial v} = X^T \left( \frac{\partial A}{\partial v} - \lambda \frac{\partial B}{\partial v} \right) X \]  \hspace{1cm} (34)

in which it is assumed that the eigenvectors are normalized such that

\[ X^T BX = 1 \]  \hspace{1cm} (35)

For eigenvector derivatives, two methods are presented by Fox and Kapoor. The first is to differentiate Eq. (30), giving a set of simultaneous equations for the eigenvalue and eigenvector derivatives. Differentiating the eigenvalue problem of Eq. (30) gives

\[ (A - \lambda B) \frac{\partial X}{\partial v} = - \left( \frac{\partial A}{\partial v} - \frac{\partial X}{\partial v} B - \lambda \frac{\partial B}{\partial v} \right) X \]  \hspace{1cm} (36)
The matrix $A - \lambda B$ is singular since $\lambda$ is an eigenvalue. The set is solvable only after algebraic manipulation, which destroys the banded nature of the equations, a point that arises later in connection with another method. The second method for eigenvector derivatives, developed by Fox and Kapoor, is to expand the derivative as a series of eigenvectors. Thus, for the $i$th eigenvector

$$\frac{\partial X_i}{\partial v} = \sum_{k=1}^{n} a_{ik} X_k \tag{37}$$

The coefficients $a_{ik}$ are obtained by substituting Eq. (37) into Eq. (36). In principle, it is necessary to use all $n$ modes in the expansion of Eq. (37). However, as with the modal method generally, it should be possible to obtain reasonable results with fewer than $n$ eigenvectors. A modification of the method of Fox and Kapoor which has exhibited faster convergence is denoted the modified modal method. This method represents the eigenvector derivative as

$$\frac{\partial X_i}{\partial v} = \left( \frac{\partial X_i}{\partial v} \right)_s + \sum_{k=1}^{n} \bar{a}_{ik} X_k \tag{38}$$

where $\left( \frac{\partial X_i}{\partial v} \right)_s$ is denoted a "psuedo static" solution which satisfies the equation

$$A \frac{\partial X_i}{\partial v} = \left( \frac{\partial \lambda}{\partial v} B - \frac{\partial A}{\partial v} + \lambda \frac{\partial B}{\partial v} \right) X_i \tag{39}$$

The coefficients $\bar{a}_{ik}$ are obtained by substituting eq. (38) into eq. (36).

Rogers and Stewart derived sensitivity formulas for eigenvalues and eigenvectors of the general problem [Eqs. (30) and (31)]. For eigenvalues the equation is

$$\frac{\partial \lambda}{\partial v} = Y^T \left( \frac{\partial A}{\partial v} - \lambda \frac{\partial B}{\partial v} \right) X \tag{40}$$

Rogers expressed the eigenvector derivatives as an expansion in terms of the eigenvectors

$$\frac{\partial X_i}{\partial v} = \sum_{k=1}^{n} a_{ik} X_k, \quad \frac{\partial Y_i}{\partial v} = \sum_{k=1}^{n} b_{ik} Y_k \tag{41}$$
The coefficients $a_{ik}$ and $b_{ik}$ are computed by substituting Eqs. (41) into an expression obtained by differentiating the eigenvalue problem and combining it with appropriate orthogonality conditions.

An alternate method for calculation of eigenvector derivatives for the symmetric problem is due to Nelson. The method of Nelson is to represent the eigenvector derivative as

$$\frac{\partial X}{\partial v} = V + cX$$

where $V$ is the solution of a reduced version of Eq. (36) obtained by deleting the $k$th row and column from $A - \lambda B$ (where $k$ is chosen to correspond to the maximum component of $X$) and setting the $k$th component of $V$ equal to zero. The multiplier $c$ is evaluated by substituting Eq. (42) into Eq. (36). This method has certain advantages over previous eigenvector derivative techniques: it requires only the eigenvalue and eigenvector for the mode being differentiated, and the equation for $V$ retains the banded character of coefficient matrix unlike the algebraic methods (e.g., Fox and Kapoor).

Repeated Eigenvalues

The sensitivity of repeated eigenvalues has been a focus of recent interest, even though the eigenvalues are not differentiable and only directional derivatives can be found. For the real symmetric case, a generalization of Nelson’s method which preserves the bandedness of the matrix was obtained by Ojalvo\(^3\) and amended by Mills-Curran\(^2\) and Daily\(^3\). These methods compute the derivatives of the $m$ eigenvectors corresponding to eigenvalues of multiplicity $m$. As stated by Dailey (see also Lancaster\(^4\)), when the eigenvalues are repeated and a design variable is perturbed, the eigenvectors “split” into as many as $m$ distinct eigenvectors. We seek the derivatives of these distinct eigenvectors which “appear” with design variable perturbation. Using Dailey’s notation, define the eigenvalue problem

$$KX = MX\Lambda,$$  \hspace{1cm} (43)

where $X$ contains the $m$ eigenvectors cited previously, and

$$\Lambda = \lambda I,$$  \hspace{1cm} (44)

where $\lambda$ is the repeated eigenvalue and $I$ is the identity matrix of order $m$. The normalization condition, Eq. (35) is now

$$X^T MX = I.$$  \hspace{1cm} (45)
The eigenvectors which appear as a result of the splitting are contained in a matrix denoted $Z$ which is related to $X$ as follows

$$Z = X$$ \hspace{1cm} (46)

where is a set of orthogonal vectors to be determined. To simplify the notation we consider a single design variable $v$, and denote derivatives with respect to that design variable by a prime. The technique for calculating $Z'$ as contained in Daily is outlined next. The vector and the derivative of the multiple eigenvalues $\Lambda'$ are obtained as solutions of the following eigenvalue problem

$$D = \Lambda',$$ \hspace{1cm} (47)

where

$$D = X^T(K' - \lambda M')X,$$ \hspace{1cm} (48)

with a normalization condition

$$T = I.$$ \hspace{1cm} (49)

Next in a manner analogous to Nelson let

$$Z' = V + ZC,$$ \hspace{1cm} (50)

where $V$ is the solution to

$$(K - \lambda M)V = (\lambda M' - K')Z + MZA',$$ \hspace{1cm} (51)

(numerically obtained by removing $m$ rows and columns from $K - \lambda M$ using the strategy described in Reference 32) and $C$ is a matrix which is obtained as the solution to the equation

$$CA' - A'C + \frac{1}{2} \Lambda'' = -Z^T(K' - \lambda M')V - Z^T(M'Z + MV)\Lambda' + \frac{1}{2} Z^T(K'' - \lambda M'')Z \equiv R.$$ \hspace{1cm} (52)

Equation (52), which requires substantial algebraic manipulations for its derivation, determines the matrix $C$ and the matrix of second derivatives of the eigenvalues $\Lambda''$. Fortunately $\Lambda''$ is diagonal and $CA' - A'C$ always has zero on the diagonal. Therefore, we can solve for the matrix $C$ separate from $\Lambda''$, and the latter matrix only needs to be calculated if it is needed for some other purpose.

13
Using Eqs. (45), (46) and (49) we have

\[ Z^T M Z = T X^T M X = I. \]  

(53)

Differentiate Eq. (53) and use Eq. (50) to obtain

\[ C + C^T = -V^T M Z - Z^T M V - Z^T M' Z \equiv Q, \]  

(54)

from which

\[ c_{ii} = \frac{1}{2} q_{ii} \]  

(55)

The non-diagonal elements of \( C \) are

\[ c_{ij} = \frac{r_{ij}}{\lambda'_i - \lambda'_j} \quad i \neq j \quad \lambda'_i \neq \lambda'_j. \]  

(56)

For the case where \( \lambda'_i = \lambda'_j \quad i \neq j \), eq. (56) may not be used. The situation here is that the eigenvalues are not "splitting" when the design variable is perturbed because the design variable is affecting both in exactly the same way. In such a case, \( Z' \) is not unique and any values of \( c_{ij} \) and \( c_{ji} \) satisfying Eq. (54) may be used. Dailey proposes the choice \( c_{ij} = c_{ji} = \frac{1}{2} q_{ij} \) whenever \( \lambda'_i = \lambda'_j \).

Before leaving the topic of derivatives associated with repeated eigenvalues, we note the limited utility of such derivatives. For example, the eigenproblem is differentiable in terms of a single parameter, but not as a function of several. This may be demonstrated by the example where the matrix

\[ A = \begin{bmatrix} 2 + y & x \\ x & 2 \end{bmatrix}. \]  

(57)

The eigenvalues of \( A \) are

\[ \lambda_{1,2} = 2 + y/2 \pm \sqrt{x^2 + y^2/4}. \]  

(58)

At \( x = y = 0 \), the eigenvalues are repeated and \( \partial \lambda / \partial x = \pm 1, \partial \lambda / \partial y = 0, 1 \). However, the eigenvalues are not differentiable as a function of both \( x \) and \( y \), that is the relation

\[ d\lambda = \frac{\partial \lambda}{\partial x} dx + \frac{\partial \lambda}{\partial y} dy \]  

(59)
does not hold. Therefore, the utility of the partial derivatives is questionable. The eigenvectors are also discontinuous at (0, 0). This can be checked by noting that at (e, 0), the eigenvectors are (1, 0) and (0, 1) and at (0, e) they are (1, 1) and (1, −1) no matter how small e is.

Sensitivity derivatives for nonlinear eigenvalue problems

In flutter and nonlinear vibration problems, we encounter eigenvalue problems where the dependence on the eigenvalue is not linear. For example, Bindolino and Mantegazza\textsuperscript{35} consider aeroelastic response problem which produces a transcendental eigenvalue problem of the form

\[ A(\lambda, v)X = 0. \] (60)

Differentiating Eq. (60) we get

\[ \frac{A\partial X}{\partial v} + \frac{\partial \lambda \partial A}{\partial v \partial \lambda} = -\frac{\partial A}{\partial v}X. \] (61)

Using the normalizing condition \( X_m = 1 \) we can solve Eq. (61) for \( \frac{\partial X}{\partial v} \) and \( \frac{\partial \lambda}{\partial v} \). Instead, it is also possible to use the adjoint method, employing the left eigenvector \( Y \) satisfying

\[ Y^T A = 0, \quad Y_m = 1 \] (62)

we obtain

\[ \frac{\partial \lambda}{\partial v} = -\frac{Y^T \frac{\partial A}{\partial v} X}{Y^T \frac{\partial A}{\partial \lambda} X}. \] (63)

A common treatment of flutter problems is to have two real parameters representing the approach of the frequency and speed as an eigenpair instead of one complex eigenvalue. For example in Murthy\textsuperscript{36}, Eq. (60) is replaced by

\[ A(M, \omega, v)X = 0, \] (64)

where the Mach number, \( M \), and the frequency, \( \omega \), are real parameters. Using this approach, differentiate Eq. (64) and premultiply by \( Y^T \) to get

\[ f_M \frac{\partial M}{\partial v} + f_\omega \frac{\partial \omega}{\partial v} = -f_v, \] (65)
where 

\[ f_M = Y^T \frac{\partial A}{\partial M} X, \quad f_\omega = Y^T \frac{\partial A}{\partial \omega} X, \quad f_v = Y^T \frac{\partial A}{\partial v} X. \]  

Multiplying Eq. (65) by \( \bar{f}_\omega \) (the complex conjugate of \( f_\omega \)) we get

\[ f_M \bar{f}_\omega \frac{\partial M}{\partial v} + |f_\omega|^2 \frac{\partial \omega}{\partial v} = -\bar{f}_\omega f_v \]  

(67)

The second term in Eq. (67) as well as \( \frac{\partial M}{\partial v} \) are real, so by taking the imaginary part of Eq. (67) we get

\[ \frac{\partial M}{\partial v} = -\frac{\text{Im}(f_\omega f_v)}{\text{Im}(f_M f_\omega)} = -\frac{\text{Im} \left[ (Y^T \frac{\partial A}{\partial M} X) \left( \bar{Y}^T \frac{\partial A}{\partial \omega} \bar{X} \right) \right]}{\text{Im} \left[ (Y^T \frac{\partial A}{\partial M} X) \left( \bar{Y}^T \frac{\partial A}{\partial \omega} \bar{X} \right) \right]} \]  

(68)

Next, multiplying Eq. (65) by \( \bar{f}_M \) and following a similar procedure gives

\[ \frac{\partial \omega}{\partial v} = -\frac{\text{Im} \left[ (Y^T \frac{\partial A}{\partial M} X) \left( \bar{Y}^T \frac{\partial A}{\partial \omega} \bar{X} \right) \right]}{\text{Im} \left[ (Y^T \frac{\partial A}{\partial M} X) \left( \bar{Y}^T \frac{\partial A}{\partial \omega} \bar{X} \right) \right]} \]  

(69)

Rudisill and Bhatia\textsuperscript{37} have a derivation of the flutter eigenpair that employs the reduced frequency and flutter speed as the eigenpair and provides also second derivatives.

It is possible to treat in a similar manner the case where the nonlinearity is in \( X \) instead of in \( \lambda \). For example, Hou et al.\textsuperscript{38} treated the nonlinear vibration problem

\[ K(X)X - \lambda MX = 0 \]  

(70)

Differentiating Eq. (70) with respect to \( v \) we obtain

\[ (J - \lambda M) \frac{\partial X}{\partial v} - \frac{\partial \lambda}{\partial v} MX = -\left[ \frac{\partial K}{\partial v} - \lambda \frac{\partial M}{\partial v} \right] X \]  

(71)

where \( J \) is the tangent stiffness matrix whose components are given as

\[ J_{ij} = K_{ij} + \sum_k \frac{\partial K}{\partial X_j} X_k \]  

(72)
Equation (71) can now be solved for eigenvector derivatives using Nelson's method. For
eigenvalue of derivatives use the left eigenvector satisfying

\[ Y^T(J - \lambda M) = 0 \]  

(73)

Premultiply Eq. (71) by \( Y^T \) to obtain

\[ \frac{\partial \lambda}{\partial v} = \frac{Y^T(\frac{\partial K}{\partial v} - \lambda \frac{\partial M}{\partial v})X}{Y^T MX} \]  

(74)

**Sensitivity of Transient Response**

**General**

The discussion of sensitivity analysis of transient structural response is usually based
on the equations of motion written as a system of second-order differential equations. However, this form obscures the similarity of structural sensitivity analysis to sensitivity analysis in other fields where first-order differential equations are employed and is also less compact than a first-order formulation. For these reasons the initial discussion will focus on a system of first-order differential equations, and the case of a second-order system will be limited to linear structural dynamics, following the general discussion. We start with a system of first-order ordinary differential equations of the form

\[ \dot{U} = F(U, t, v) \]

\[ U(0) = U_o \]  

(74)

where \( U \) is the response, \( F \) a vector of functions, \( t \) time, and \( v \) a typical design parameter; a dot denotes differentiation with respect to time. In many structural applications the left-hand side of Eqs. (56) is \( A \dot{U} \), where \( A \) is a matrix, and the methods discussed below are also applicable to that more general form.

**Direct Method**

The direct method of obtaining sensitivity derivatives is based on differentiating
Eqs. (74) to obtain

\[
\frac{dU}{dv} - J \frac{dU}{dv} = \frac{\partial F}{\partial v}
\]

(75)

\[
\frac{dU}{dv}(0) = 0
\]

where the Jacobian \( J \) is \( \partial f/\partial U \). Note that Eqs. (75) are a system of linear differential equations, even if the original system, Eqs. (79), is nonlinear. Often derivatives of the entire vector \( U \) are not required. Instead it is necessary to obtain the derivatives of a function of \( U \) of the form.

\[
g(U, v) = \int_0^{t_f} p(U, t, v) dt
\]

(76)

where \( p \) is a functional representation of a time-dependent constraint and \( t_f \) is a final time for the response calculation. The direct approach obtains \( dg/dv \) as

\[
\frac{dg}{dv} = \int_0^{t_f} \left[ \frac{\partial p}{\partial v} + \left( \frac{\partial p}{\partial U} \right)^T \frac{dU}{dv} \right] dt
\]

(77)

where \( dU/dv \) is calculated in Eqs. (75).

**Green's Function Method**

Equation (75) have to be solved once for each design variable and are costly when the number of design variables is large. When the number of design variables is larger than the dimensionality of \( U \), then the Green's function approach is more efficient than the direct approach. An application of this approach is sensitivity analysis of transient structural response, when the response is computed using reduction techniques such as modal analysis. The sensitivity derivatives \( dU/dv \) is written as

\[
\frac{dU}{dv}(t) = \int_0^t K(t, \tau) \frac{\partial F}{\partial v}(\tau) d\tau
\]

(78)
where the Green’s function $K$ satisfies (recall that the dot denotes $d/dt$)

$$
K(t, \tau) = 0, \quad t < \tau
$$

$$
K(\tau, \tau) = I
$$

$$
\dot{K}(t, \tau) - J(t)K(t, \tau) = 0, \quad t > \tau
$$

(79)

The efficiency of the Green’s function approach is partly governed by the method used to integrate Eqs. (79). A large amount of work on the efficient implementation of the Green’s function approach has been performed by Rabitz and co-workers.\textsuperscript{39}

**Adjoint Variable Method**

Further improvements in efficiency may be possible if less information is needed. If instead of the derivatives of the entire vector $U$, only those of a few functionals [e.g., Eq. (76)] are required, then an adjoint variable method is called for. The adjoint variable approach solves first for the adjoint vector $\lambda$ from the differential equation

$$
\dot{\lambda} + J^T\lambda = \frac{\partial p}{\partial U}
$$

$$
\lambda(t_f) = 0
$$

(80)

It is shown by Haftka, Gürdal and Kamat\textsuperscript{40} that

$$
\frac{dg}{dv} = \int_0^{t_f} \left( \frac{\partial p}{\partial v} - \lambda^T \frac{\partial F}{\partial v} \right) dt
$$

(81)

Equations (80) are a set of linear differential equations that is integrated backward from $t_f$ to zero. As in the steady-state case, the adjoint variable approach is preferred over the direct approach when the number of functionals is less than the number of design variables.

**Finite Difference Method**

For sensitivity analysis of static response, the finite difference approach is almost always inferior to analytical methods. For the calculation of derivatives of transient response,
this is not always the case. When explicit methods are used for integrating the differential equations, the linearity of the sensitivity equations does not constitute a computational advantage. Therefore, for the case of explicit integration, the finite difference approach is often computationally superior to the direct method (see Ref. 41). When implicit integration techniques are used, the finite difference approach is less attractive computationally but remains easier to implement than the direct approach.

Linear Structural Dynamics

For the case of linear structural dynamics it may be advantageous to retain the second-order equations of motion rather than reduce them to a set of first-order equations. It is also common to use modal reduction for this case. In this section we discuss the application of the direct and adjoint methods to this special case. The equations of motion are written as

\[ M\ddot{U} + C\dot{U} + KU = F(t) \]  

Most often the problem is reduced in size by expressing \( U \) in terms of \( m \) basis functions \( \Phi^i, i = 1, \ldots, m \) where \( m \) is usually much less than the number of degrees of freedom of the original system Eq.(82)

\[ U = \Phi Q \]  

where \( \Phi \) is a matrix with \( \Phi^i \) as columns. Then a reduced set of equations can be written as

\[ \bar{M}\ddot{Q} + \bar{C}\dot{Q} + \bar{K}Q = \bar{F}(t) \]  

where

\[ \bar{M} = \Phi^T M \Phi, \quad \bar{C} = \Phi^T C \Phi, \quad \bar{K} = \Phi^T K \Phi, \quad \bar{F} = \Phi^T F \]  

When the basis functions are the first \( m \) natural vibration modes of the structure scaled to unit modal masses \( \Phi \) satisfies the equation

\[ K\Phi - M\Phi\Omega^2 = 0 \]  

where \( \Omega \) is a diagonal matrix with the \( i \)th natural frequency \( \omega_i \) in the \( i \)th row. In that case \( \bar{K} = \Omega^2 \) and \( \bar{M} = I \) are diagonal matrices. For special forms of damping, the damping matrix \( \bar{C} \) is also diagonal so that the system Eq. (84) is uncoupled. After \( Q \) is calculated from Eq. (84) we can use Eq. (83) to calculate \( F \). This method is known as the mode-displacement method.

When the load \( F \) has spatial discontinuities the convergence of the modal approximation, Eq. (83), can be very slow. The convergence can be dramatically accelerated by
using the mode acceleration method. The mode acceleration method can be derived by rewriting Eq. (82) as

\[ U = K^{-1}F - K^{-1}C\ddot{U} - K^{-1}M\dddot{U} \]  

(87)

The first term in Eq. (87) is called the quasi-static solution because it represents the response of the structure if the loads are applied very slowly. The second term and third terms are approximated in terms of the modal solution. It can be shown (e.g., Greene42) that \( K^{-1} \) can be approximated as

\[ K^{-1} = \Phi \Omega^{-2} \Phi^T \]  

(88)

Using this approximation for the second and third terms of Eq. (87) we get

\[ U \approx K^{-1}F - \Phi \Omega^{-2}C\dot{Q} - \Phi \Omega^{-2}\ddot{Q} \]  

(89)

This approximation is exact when \( \Phi \) contains the full set of vibration modes. Note that \( \dot{Q} \) and \( \ddot{Q} \) in Eq. (89) are obtained from the mode-displacement solution, Eq. (84). Therefore, there is no difference in velocities and accelerations between the mode-displacement and the mode acceleration.

In considering the calculation of sensitivities we treat first the mode-displacement method. The direct method of calculating the response sensitivity is obtained by differentiating Eq. (84) to obtain

\[ \tilde{M} \frac{d\dot{Q}}{dv} + \tilde{C} \frac{d\ddot{Q}}{dv} + \tilde{K} \frac{d\dddot{Q}}{dv} = R \]  

(90)

where

\[ R = \frac{d\dddot{F}}{dv} - \frac{d\tilde{M}}{dv} \ddot{Q} - \frac{d\tilde{C}}{dv} \dot{Q} - \frac{d\tilde{K}}{dv} Q \]  

(91)

The first step in forming this equation is the calculation of the derivatives of \( \dddot{F} \), \( \tilde{M} \), \( \tilde{C} \), and \( \tilde{K} \) with respect to \( v \). Differentiation of \( \tilde{K} \) yields

\[ \frac{d\tilde{K}}{dv} = \Phi^T \frac{d\Phi}{dv} + \frac{d\Phi^T}{dv} K \Phi + \Phi^T K \frac{d\Phi}{dv} \]  

(92)

with similar expressions for the derivatives of \( \tilde{M} \), \( \tilde{C} \), and \( \dddot{F} \). The calculation is simplified considerably by using a fixed set of basis functions \( \Phi \) or neglecting the effect of the change in the modes. In many cases the error associated with neglecting the effect of changing
modes is small. When this error is unacceptable we have to face the costly calculation of the
derivatives of the modes needed for calculating the derivatives of the reduced matrices, such
as Eq. Fortunately it was found by Greene\textsuperscript{42} that the cost of calculating the derivatives of
the modes can be substantially reduced by using the modified modal method\textsuperscript{26}, keeping
only the first term in this equation. This approximation to the derivatives of the modes
may not always be accurate, but it appears to be sufficient for calculating the sensitivity
of the dynamic response.

For the adjoint method we consider a function in the form of Eq. (76)

$$g(Q, v) = \int_0^{t_f} p(Q, v, t) dt$$

so that

$$\frac{dg}{dv} = \int_0^{t_f} \left( \frac{\partial p}{\partial v} + \frac{\partial p}{\partial Q} \frac{dQ}{dv} \right) dt$$

To avoid the calculation of $dQ/dv$ we use an adjoint vector $\lambda$, and start by multiplying
Eq. (90) by $\lambda^T$ and integrating

$$\int_0^{t_f} \lambda^T (\ddot{\bar{M}} \frac{d\bar{Q}}{dv} + \bar{C} \frac{d\bar{Q}}{dv} + \bar{K} \frac{dQ}{dv}) dt = \int_0^{t_f} \lambda^T Rd\lambda$$

Integrating by parts we get

$$\lambda^T \ddot{\bar{M}} \frac{d\bar{Q}}{dv} \bigg|_0^{t_f} - \dot{\lambda}^T \ddot{\bar{M}} \frac{d\bar{Q}}{dv} \bigg|_0^{t_f} + \lambda^T \bar{C} \frac{d\bar{Q}}{dv} \bigg|_0^{t_f} + \int_0^{t_f} (\dot{\lambda}^T \ddot{\bar{M}} - \dot{\lambda}^T \bar{C} + \lambda^T \bar{K}) \frac{dQ}{dv} dt = \int_0^{t_f} \lambda^T Rd\lambda$$

Assuming that the initial conditions do not depend on the design variable $v$, Eq. (96)
suggests the following definition for $\lambda$

$$\ddot{\bar{M}} \ddot{\lambda} - \ddot{\bar{C}} \lambda + \ddot{\bar{K}} \lambda = (\frac{\partial p}{\partial Q})^T, \quad \lambda(t_f) = \dot{\lambda}(t_f) = 0$$
and then Eq. (94) becomes

\[
\frac{dg}{dv} = \int_0^t \left( \frac{\partial p}{\partial v} - \lambda^T R \right) dt
\]  

(98)

For the mode-acceleration method we consider only the direct method. We start by differentiating Eq. (82) and rearranging it as

\[
\frac{dU}{dv} = K^{-1} \left[ \frac{dF}{dv} - \frac{dK}{dv} [U - C \frac{dU}{dv} - C \frac{d\dot{U}}{dv} - \frac{dM}{dv} \frac{\ddot{U}}{dv}] \right]
\]  

(99)

Next we use Eq. (89) to approximate the second term, and the modal expansion Eq. (83) to approximate the other terms to get

\[
\frac{dU}{dv} \approx K^{-1} \left[ \frac{dF}{dv} - \frac{dK}{dv} \left[ K^{-1} F - \Phi \Omega^{-2} \ddot{Q} - \Phi \Omega^{-2} \dddot{Q} \right] - C \Phi \frac{d\ddot{Q}}{dv} - M \Phi \frac{d\dddot{Q}}{dv} - \frac{dM}{dv} \frac{\dddot{Q}}{dv} \right]
\]  

(100)

Finally we use the modal approximation to \( K^{-1} \), Eq. (88) to obtain

\[
\frac{dU}{dv} \approx K^{-1} \left[ \frac{dF}{dv} - \frac{dK}{dv} K^{-1} F \right] + \Phi \Omega^{-2} \Phi^T \left[ \frac{dK}{dv} \Phi \Omega^{-2} \ddot{Q} - \frac{dC}{dv} \Phi \dddot{Q} - C \Phi \frac{d\dddot{Q}}{dv} \right] + K^{-1} \left[ \frac{dK}{dv} \Phi \Omega^{-2} - \frac{dM}{dv} \Phi \right] \dddot{Q} - \Phi \Omega^{-2} \frac{d\dddot{Q}}{dv}
\]  

(101)

Note that the calculation involves the solution of Eqs. (84) and (90) for \( Q \) and \( dQ/dv \), followed by Eq. (101) for retrieving the \( dU/dv \). Additional details can be found in\textsuperscript{42}. 

23
References


This paper presents a descriptive review of the most recent methods for performing sensitivity analysis of the structural behavior of discretely-modeled systems. The methods are generally but not exclusively aimed at finite element modeled structures. Topics included in the paper in order of presentation are: selection of finite difference step sizes; special consideration for finite difference sensitivity of iteratively-solved response problems; first and second derivatives of static structural response; sensitivity of stresses; nonlinear static response sensitivity; eigenvalue and eigenvector sensitivities for both distinct and repeated eigenvalues; and sensitivity of transient response for both linear and nonlinear structural response.