AIAA 81-0548R

Sensitivity of Optimum Solutions of Problem Parameters

Jaroslaw Sobieszczanski-Sobieski*
NASA Langley Research Center, Hampton, Va.
Jean-Francois Barthelemy†
Virginia Polytechnic Institute and State University, Blacksburg, Va.
and
Kathleen M. Riley‡
Kentron International, Inc., Hampton, Va.

Solution of the optimum sensitivity problem yields the values of derivatives of the optimal objective function and design variables with respect to those physical quantities which were kept constant as problem parameters during optimization. Examples of these sensitivity derivatives might include derivatives of cross-sectional area and structural mass with respect to allowable stress and derivatives of fuel consumed and wing aspect ratio with respect to aircraft range. Derivation of the sensitivity equations that yield the sensitivity derivatives directly, which avoids the costly and inaccurate "perturb-and-reoptimize" approach, is discussed and solvability of the equations is examined. The equations apply to optimum solutions obtained by direct search methods as well as those generated by procedures of the sequential unconstrained minimization technique (SUMT) class. Applications are discussed for the use of the sensitivity derivatives in extrapolation of the optimal objective function and design variable values for incremented parameters, optimization with multiple objectives, and decomposition of large optimization problems. Several aspects of these applications and verification of the sensitivity equations are presented through numerical examples.

Nomenclature

a	= weighting factor
A	= cross-sectional area
b	= behavior variable
C	= convergence tolerance governing termination of optimization
f	= function in general mathematical sense
F	= objective function
g	= vector of m constraints
g_j	= jth constraint, assumed to be active unless noted otherwise; constraint is violated when $g_j > 0$
m	= number of constraints active at constrained minimum
n	= number of design variables in optimization problem
P	= penalty term in penalty function ϕ, and force
p_k	= kth parameter; in differentiation with respect to p_k, k is omitted
r	= drawdown factor in SUMT
W	= weight
X, x_i	= vector of design variables and ith design variable, respectively
ϵ	= constraint violation tolerance (typically a small number, e.g., 0.025)
λ	= vector of m Lagrange multipliers
λ_j	= jth multiplier corresponding to jth active constraint
ϕ	= penalty function

Subscripts and Superscripts

a	= allowable value
e	= extrapolated value
h	= number of objectives in a multiobjective optimization
i	= ith design variable, ith weighting factor
j	= jth constraint, jth weighting factor
k	= kth parameter
l	= lower bound
0	= initial value in extrapolation
p	= prescribed quantity
q	= refers to qth design variable
u	= upper bound
$(\bar{\ })$	= quantities at optimum

Differential Notation

$f^{(i)}$	$\equiv \partial f/\partial x_i$; $f^{(i,q)} \equiv \partial^2 f/\partial x_i \partial x_q$
f'	$\equiv \partial f/\partial p$ assuming constant x_i values
$f'^{(i)}$	$\equiv \partial^2 f/\partial x_i \partial p$

In this notation

$\dot{g}$	$\equiv [\dot{g}_1, \dot{g}_2, \ldots \dot{g}_j, \ldots \dot{g}_m]$, $n \times m$ matrix
$\dot{g}_j$	$\equiv \{\dot{g}_j^{(1)}, \dot{g}_j^{(2)}, \ldots \dot{g}_j^{(i)}, \ldots \dot{g}_j^{(n)}\}$, column vector of length n
$\dot{g}^{(i)}$	$\equiv \{\dot{g}_1^{(i)}, \dot{g}_2^{(i)}, \ldots, \dot{g}_j^{(i)}, \ldots \dot{g}_m^{(i)}\}$, row vector of length m
$\ddot{g}^{(i,q)}$	$\equiv \{\ddot{g}_1^{(i,q)}, \ddot{g}_2^{(i,q)}, \ldots \ddot{g}_m^{(i,q)}\}$, row vector of length m

Presented as Paper 81-0548 at the AIAA/ASME/ASCE/AHS 22nd Structures, Structural Dynamics and Materials Conference, Atlanta, Ga., April 6-8, 1981; submitted June 1, 1981; revision received Jan. 4, 1982. This paper is declared a work of the U.S. Government and therefore is in the public domain.

*Head, Multidisciplinary Analysis and Optimization Branch, Loads and Aeroelasticity Division. Associate Fellow AIAA.

†Graduate Student, Department of Aerospace and Ocean Engineering. Student Member AIAA.

‡Scientific Programmer, Hampton Technical Center.

Introduction

NONLINEAR mathematical programming has become well established as a tool for defining optimal engineering designs as local constrained minima. A typical constrained minimization problem entails a group of physical quantities which are used as design variables and a group of constant quantities termed parameters of the problem. It is of obvious interest to know, when the optimization is completed, the sensitivity of the constrained minimum to the parameters of the problem. Mathematically, this requires the determination of the partial derivatives of the objective function and design variables with respect to the parameters of interest. These derivatives are referred to as sensitivity derivatives. For example, in structural optimization, it would be useful to determine the effect on optimal structural mass and cross-sectional dimensions of changes in allowable stress or displacements. In an aircraft configuration optimization, the information of interest would be the sensitivity of optimal block fuel consumption and wing aspect ratio and area to variations in required range and payload.

Generation of sensitivity derivatives by finite-difference approximations requires reoptimization of the problem with incremented values of the parameters. This is a costly procedure burdened with the difficulty of assessing numerical errors. A preferable approach is to obtain the sensitivity derivatives directly from an appropriate set of equations. This approach, known as optimum sensitivity analysis, became well established as a routine tool in linear programming.[1] In contrast, sensitivity analysis in nonlinear mathematical programming is still at an early stage of development, and its incorporation into engineering practice is yet to be accomplished. The relatively limited literature available on the subject, primarily in the discipline of operating research, is well represented by Refs. 2-5. Specifically, Ref. 2 gives a solution to the problem of finding the increments of the optimum objective function and variables caused by simultaneous small changes of the problem parameters, and contains references to other relevant early works. Reference 3 develops a solution for sensitivity in the context of a sequential unconstrained minimization technique (SUMT) and a particular form of a penalty function. Some computational experience with that solution, including an application in resource management, is reported in Ref. 4. With the exception of a simple torsion bar example in Ref. 5, the authors are aware of no reported research on optimum sensitivity analysis applied to optimization of structures and other engineering systems.

The objectives of this paper are to show how the equations capable of yielding the sensitivity derivatives (the sensitivity equations) can be obtained for a constrained optimum regardless of the type of optimization algorithm that was used to arrive at the optimum point; to review the solvability of the sensitivity equations; and to report on applications in structural optimization. Lagrange multiplier equations and the extremum conditions of a penalty function in general, and the interior and exterior forms of that function in particular, are included as alternative bases for derivation of the sensitivity equations. Numerical examples selected from Ref. 6 verify the algorithms and illustrate evaluation of the sensitivities of optimal structures to parameters such as load, allowable stress, and overall geometrical shape and the influence of individual conflicting objectives (e.g., weight and cost) in multiobjective optimization. An additional application of the sensitivity analysis, to provide a basis for a formal decomposition in hierarchical optimization problems, is also discussed.

Sensitivity Equations

There are at least two ways of deriving equations for the unknown sensitivities. One way is to start from the Lagrange multiplier equations of the constrained minimum; the other begins with the extremum conditions of a penalty function. In either case the same general functional relationships are recognized, namely, the objective function

$$F(x_1,x_2,\ldots x_n,\ p_1,p_2,\ldots p_k) \tag{1a}$$

and constraints active at the optimum point

$$g_j(x_1,x_2,\ldots x_n,\ p_1,p_2,\ldots p_k) \tag{1b}$$

and implicitly

$$x_i(p_1,p_2,\ldots p_k) \tag{1c}$$

There is no need to distinguish among the equality and inequality constraints in the original formulation of the optimization problem being analyzed for sensitivity, because all of the constraints active at the constrained optimum point may be regarded as equality constraints.

Sensitivity Equations Derived from the Lagrange Multiplier Equations

The familiar Lagrange multiplier equations satisfied at a constrained minimum are

$$\dot{F}^{(i)} + \dot{g}^{(i)}\lambda = 0, \qquad i = 1 \rightarrow n \tag{2a}$$

$$g_j = 0, \qquad j = 1 \rightarrow m \tag{2b}$$

The equations may be differentiated with respect to the parameter p_k using the chain-differentiation rule for composite functions along with the functional relationships in Eq. (1). The result of the differentiation is

$$\dot{F}'^{(i)} + \sum_{q=1}^{n} \ddot{F}^{(i,q)} x_q' + \sum_{j=1}^{m}\left[\dot{g}_j'^{(i)} + \left(\sum_{q=1}^{n} \ddot{g}_j^{(i,q)} x_q'\right)\right]\lambda_j + \sum_{j=1}^{m} \dot{g}_j^{(i)}\lambda_j' = 0, \qquad i = 1 \rightarrow n \tag{3a}$$

$$g_j' + \sum_{i=1}^{n} \dot{g}_j^{(i)} x_i' = 0, \qquad j = 1 \rightarrow m \tag{3b}$$

Equations (3) can be converted to a uniform matrix notation by collecting terms and using an auxiliary matrix Z

$$\underset{(n+m)\times(n+m)}{\begin{bmatrix} \underset{n\times n}{[\ddot{F}+Z]} & \underset{n\times m}{[\dot{g}]} \\ \underset{m\times n}{[\dot{g}]^T} & \underset{m\times m}{[0]} \end{bmatrix}} \underset{(n+m)\times 1}{\begin{Bmatrix} \bar{X}' \\ \lambda' \end{Bmatrix}} + \underset{(n+m)\times 1}{\begin{Bmatrix} \underset{n\times 1}{\{\dot{F}'\}} + \underset{n\times m}{[\dot{g}']}\underset{m\times 1}{\{\lambda\}} \\ \underset{m\times 1}{\{g'\}} \end{Bmatrix}} = 0 \tag{4a}$$

where the dimensions of vectors and matrices are inscribed for clarity and where Z is defined as a square, $n \times n$, matrix whose i,q element is

$$Z^{(i,q)} = \sum_{j=1}^{m} \frac{\partial^2 g_j}{\partial x_i \partial x_q}\lambda_j \tag{4b}$$

Equations (4), whose terms are evaluated at the constrained minimum point, constitute a set of $n+m$ simultaneous linear algebraical equations for unknown derivatives of the optimum solution $\bar{X}'$ (n elements) and λ' (m elements), the latter being auxiliary quantities.

Once the sensitivity derivatives $\bar{x}_i'$ are obtained, the sensitivity derivative of the objective function is determined as total derivative of the composite function F

$$\mathrm{d}\bar{F}/\mathrm{d}p = \bar{F}' + \sum_i \dot{\bar{F}}^{(i)} \bar{x}_i' \tag{5}$$

Solution of the Sensitivity Equations

The ways to obtain a solution of $\bar{X}'$ (and λ') from Eqs. (4) may be categorized in a number of cases.

Case 1: All Submatrices in Eqs. (4) Exist

This basic case corresponds to constrained minimum defined by a nonlinear objective function and nonlinear constraint functions. The Lagrange multipliers are available as a by-product of the optimization solution. Derivative vectors $\dot{g}_j$ must be linearly independent in order for the matrix of coefficients in Eqs. (4) to be nonsingular. A Gram matrix test,[7] in addition to physical insight, may be used to identify redundant constraints to be eliminated. No ill-conditioning difficulties were observed when solving Eqs. (4) for test problems using a standard Gaussian elimination solution algorithm. A particular case of $m=n$ causes $\bar{X}'$ in Eqs. (4) to decouple from λ' (see discussion of case 4).

Case 2: Multipliers λ_i Not Available from the Optimum Solution

The Lagrange multipliers λ_i must be known in order to construct the matrix of coefficients [they enter matrix Z and the vector of free terms in Eqs. (4)]. If the optimum being analyzed for sensitivity is obtained by a method that does not yield the λ values as part of the solution, these values have to be computed as part of the input to Eqs. (4) [for an exception see Eq. (8) and the discussion of case 4]. One well-known relationship[8] that may be used for computing the λ values is

$$\lambda = -[\dot{g}^T\dot{g}]^{-1}\dot{g}^T\dot{F} \tag{6}$$

where the matrix $\dot{g}$ contains the constraints active at the optimum. It is important again that only the linearly independent constraints among the active set be included in Eq. (6), as mentioned in the discussion of case 1.

Case 3: Linearity of Constraints or Objective Function

The linearity of constraints at the optimum eliminates the matrix Z, while the linearity of the objective function renders $\ddot{F}=0$. However, if the linearity of the constraints and objective function do not occur simultaneously, the term $\ddot{F}+Z$ does not vanish and the unknowns $\bar{X}'$ and λ' remain coupled in Eqs. (4).

This case also includes the constrained minima in which some of the active constraints are side constraints. For the purposes of sensitivity analysis, such constraints should be reformulated as inequality constraints

$$g_j = 1 - x_i/x_{il} \leq 0$$

or

$$g_j = x_i/x_{iu} - 1 \leq 0 \tag{7}$$

for lower and upper bounds, respectively. This formulation leads to derivatives with respect to x_i, x_{il}, and x_{iu} and also permits sensitivity analysis with respect to the side constraint parameters such as x_{il} and x_{iu}.

Case 4: Linear Constraints and Linear Objective Function

Under this condition $\ddot{F}=0$ and $Z=0$, hence the term $\ddot{F}+Z$ vanishes and the $\bar{X}'$ and λ' vectors decouple in Eqs. (4) so that

$$\underset{m\times n}{[\dot{g}^T]}\underset{n\times 1}{\{\bar{X}'\}} + \underset{m\times 1}{\{g'\}} = \{0\} \tag{8}$$

which is an equation that does not contain the Lagrange multipliers λ. The solvability of Eq. (8) depends on the dimensions m and n and requires consideration of the following subcases that are likely to occur in practice.

Subcase 1. The $\dot{g}^T$ is a square matrix, $m=n$, and there are no null rows and columns. Consequently, Eq. (8) is determined so that a unique solution can be obtained for $\bar{X}'$. Typically, this subcase occurs when a nonlinear mathematical programming problem is solved by a sequence of steps, each step consisting of finding a constrained minimum of the problem that is locally linearized and subjected to move limits expressed in form of Eqs. (7). Such a minimum falls on a full vertex of the linearized feasible domain, hence $m=n$.

Subcase 2. The $\dot{g}^T$ is a rectangular matrix with $m<n$ rendering Eq. (8) underdetermined. This subcase arises when a constrained minimum in a nonlinear problem is defined, at least in part, by the tangency of the constraint boundary and constant objective function hypersurfaces whose curvatures, represented by the second derivatives with respect to the design variables, are not available among the optimization results and are too costly to compute.

Sensitivity Equations Derived from Extremum Conditions of a Penalty Function

In a broad class of methods known as sequential unconstrained minimization techniques (SUMT), the objective function is augmented by a penalty term containing the constraints and an approximation to the constrained minimum is determined asymptotically by generating a series of unconstrained minima of a penalty function.[2,8] Sensitivity equations analogous to Eqs. (4) may be obtained for a SUMT-determined constrained minimum by differentiating the extremum conditions of a penalty function with respect to a parameter.

There are many formulations of the penalty functions currently in use, for example, interior, exterior,[8] and quadratic extended[9] penalty function formulations. Therefore, a penalty function in its most general form is expressed as

$$\phi = F + rP \tag{9}$$

The penalty term P is a function of the constraints

$$P(g_j), \qquad j = 1 \to m \tag{10}$$

The extremum conditions for ϕ are

$$\dot{\phi}^{(i)} = \dot{F}^{(i)} + r\sum_{j=1}^{m} \frac{\partial P}{\partial g_j}\dot{g}_j^{(i)} = 0, \qquad i = 1 \to n \tag{11}$$

Differentiation of Eq. (11) with respect to parameter p_k yields

$$\sum_{q=1}^{n}\left\{\left[\ddot{F}^{(i,q)} + r\sum_{j=1}^{m}\left(\frac{\partial^2 P}{\partial^2 g_j}\dot{g}_j^{(i)}\dot{g}_j^{(q)} + \frac{\partial P}{\partial g_j}\ddot{g}_j^{(i,q)}\right)\right]\bar{x}'_q\right\}$$
$$+\dot{F}'^{(i)} + r\sum_{j=1}^{m}\left(\frac{\partial^2 P}{\partial g_j^2}g'_j\dot{g}_j^{(i)} + \frac{\partial P}{\partial g_j}\dot{g}_j'^{(i)}\right) = 0, \quad i = 1 \to n \tag{12}$$

Equation (12) represents a set of n simultaneous linear equations for n unknown values $\bar{x}'_q$.

For a specific case of an interior penalty function in a frequently used form

$$P = -\sum_{j=1}^{m} 1/g_j \tag{13}$$

Eq. (11) becomes

$$\dot{\phi}^{(i)} = \dot{F}^{(i)} + r\left(\sum_{j=1}^{m} g_j^{-2}\dot{g}_j^{(i)}\right) = 0 \tag{14}$$

and Eq. (12) yields

$$\underset{n\times n}{[M]}\underset{n\times 1}{\{\bar{X}'\}} + \underset{n\times 1}{\{R\}} = 0 \tag{15}$$

where

$$M_{iq} = \ddot{F}^{(i,q)} - 2r\sum_{j=1}^{m} g_j^{-3}\dot{g}_j^{(i)}\dot{g}_j^{(q)} + r\sum_{j=1}^{m} g_j^{-2}\ddot{g}_j^{(i,q)} \tag{15a}$$

$$R_i = \dot{F}'^{(i)} - 2r\sum_{j=1}^{m} g_j^{-3}g'_j\dot{g}_j^{(i)} + r\sum_{j=1}^{m} g_j^{-2}\dot{g}_j'^{(i)} \tag{15b}$$

At this point one may be tempted to take advantage of the fact that $g_j^{-2} \ll g_j^{-3}$ for active constraints which are very small by definition, and to neglect terms with g_j^{-2} (second derivatives would be eliminated from the calculations) in Eq. (15)—a simplification that has been introduced into optimization practice by Ref. 10. However, such a simplification cannot be made in this case because it renders Eq. (15) singular.

Turning now to an exterior penalty function whose penalty term is

$$P=\sum_{j=1}^{m}(\langle g_j\rangle^2;\ \langle g_j\rangle=\begin{cases} g_j, \text{ if } g_j>0 \\ 0., \text{ if } g_j\leq 0\end{cases} \tag{16}$$

one reduces Eq. (12) again to the form of Eq. (15) in which M and R become

$$M_{iq}=\ddot{F}^{(i,q)}+2r\sum_{j=1}^{m}(\dot{g}_j^{(i)}\dot{g}_j^{(q)}+g_j\ddot{g}_j^{(i,q)}) \tag{17a}$$

$$R_i=\dot{F}'^{(i)}+2r\sum_{j=1}^{m}(g_j'\dot{g}_j^{(i)}+g_j\dot{g}_j'^{(i)}) \tag{17b}$$

and all constraint functions are in the $\langle g_j\rangle$ form defined by Eq. (16).

Other forms of penalty functions can be treated in the same manner (e.g., Ref. 4).

Numerical values of the objective and constraint functions, their derivatives, and in a general case, the value of r are necessary for generating sensitivity equations such as Eq. (15). The value of r may not be available for the particular case of a constrained minimum being analyzed for sensitivity. However, it can be computed for the given set of constrained minimum point coordinates from Eq. (11) if that point's location is slightly off the constraint boundary in the feasible region (a typical outcome of a SUMT using an interior penalty function). If the location is slightly off into the unfeasible region (a typical result of the use of an exterior penalty function) the corresponding extremum condition equations for an exterior penalty function can be used to calculate r. Any equation arbitrarily chosen among n equations involved may be used to calculate the single unknown r, but the accuracy of this calculation will obviously benefit from choosing an r that minimizes the sum of the squares of residuals of Eq. (11), thus solving these equations in a least squares sense. The algebraic structure of the sensitivity equations for the penalty function makes their solution independent of the constant r in cases for which $\ddot{F}^{(i,q)}\equiv 0$ and $\dot{F}'^{(i)}\equiv 0$. These cases include, for example, sensitivity of minimum structural mass with respect to parameters other than the specific mass and overall geometrical dimensions.

In both Eqs. (4) and (15), only the right-side vector contains the mixed derivatives $(\cdot\,')$ that depend on the choice of parameter p with respect to which the sensitivity analysis is to be carried out. This property obviously reduces the amount of numerical labor required in problems with many parameters.

Cross Applications of Eqs. (4) and (12)

Although the origin of Eqs. (4) seems to suggest that it is appropriate for a sensitivity analysis of constrained minima obtained by one of the direct minimization methods (i.e., usable-feasible directions technique) and, by the same token, Eq. (12) seems to apply naturally in sensitivity analysis of constrained minima found by a SUMT approach, these equations can be cross applied. In such cross applications, Eq. (6) can be used to determine the λ values and Eq. (11) to compute the r value, so that one can apply Eqs. (4) (and the corresponding solution subcase equations) to analyze the sensitivity of the SUMT-obtained constrained minima and, conversely, Eq. (12) can be used when a constrained minimum is determined by a direct-search algorithm.

Potential Applications

The primary application that inspired the method discussed herein is the enhancement of the results of optimization with sensitivity data, thus increasing, at a relatively small additional cost, the information about the object being optimized, including a possibility of extrapolation with respect to incremented parameters. However, the method appears capable of additional applications which are described in this section.

Extrapolation

The sensitivity derivatives provide an estimate of the change in the objective function and design variables corresponding to a small change of a parameter. Using the first two terms of the Taylor series expanded about the optimum design point $\bar{x}_0$

$$F_e=\bar{F}_0+(\mathrm{d}\bar{F}/\mathrm{d}p)\Delta p \tag{18a}$$

$$x_{ei}=\bar{x}_{0i}+\bar{x}_i'\Delta p; \qquad i=1\rightarrow n \tag{18b}$$

Alternatively, for F computed as $\bar{F}_o=f(\bar{x}_{oi})$ at optimum:

$$F_e=f(x_{ei}) \tag{18c}$$

which suggests recomputing F using the x_{ei} estimates obtained from Eq. (18b).

Optimization with Multiple Objectives

Sensitivity analysis may be used to evaluate the influence of each of the multiple objectives on a constrained minimum. In a multiple objective optimization one may use a composite objective function in the form of a weighted sum of the single objectives

$$F=a_1F_1+a_2F_2+\ldots a_iF_i+\ldots a_hF_h \tag{19}$$

where the weighting factors a_i represent the relative importance of the individual objectives F_i.

Since each weighting factor a_i is a problem parameter, the sensitivities of the optimum with respect to the a_i factors can be obtained to determine the influence these factors have on the optimum and the trends associated with changes to the a_i values. It is conceivable that in well-behaved problems it may be possible to use the derivatives with respect to a_i for extrapolation of $\bar{F}$ and $\bar{x}_i$ values to estimate their magnitudes for $a_i=1$, and $a_j=0$ for $j\neq i$. In effect, this would provide estimates of single-objective optimizations for each of the h objectives F_i, at a computational cost not much larger than the cost of a single optimization.

Predicting a Change in Constraint Status

Considering the active constraint g_j and the corresponding λ_j value, a simple extrapolation of λ_j to zero

$$\lambda_{ej}=\lambda_{0j}+\lambda_j'\Delta p=0 \tag{20a}$$

yields an estimate

$$\Delta p=-\lambda_{0j}/\lambda_j' \tag{20b}$$

for the increment Δp of parameter p sufficient to make that active and nonredundant constraint an active but redundant one. Analogously, one may estimate an increment Δp needed for an inactive constraint, whose value g_j is less than $-\epsilon$ at the constrained optimum, to become active

$$(g_j)_{\text{new}}=-\epsilon=(g_j)_{\text{old}}+\Big(g_j'+\sum_{i=1}^{n}\dot{g}_j^{(i)}\bar{x}_i'\Big)\Delta p \tag{21}$$

where the $\bar{x}_i'$ values are available from the sensitivity analysis, and the values of g_j' and $\dot{g}_j^{(i)}$ would have to be computed additionally for the previously inactive constraint g_j.

Decomposition of a Large Optimization Problem into Several Subproblems

Another application is a decomposition of large multivariable optimization problems into a number of smaller subproblems. Suppose that the vector (X) of n design variables in an optimization problem is partitioned into X_A and X_B parts of n_A and n_B respective lengths and that an optimization is performed with X_A variables, while holding elements of (X_B) fixed as parameters of the problem. Applying the optimum sensitivity analysis, the partial derivatives of the objective function and elements of (X_A) with respect to elements of (X_B) can be obtained. These derivatives, used in a first-order Taylor series, permit expression of the objective function and each element of (X_A) as approximate linear functions of (X_B), thus eliminating (X_A) from the problem in the vicinity of the design point where the derivatives were evaluated. This capability to eliminate a group of variables may be used as means for formal decomposition of a large optimization problem into a hierarchy of subproblems coupled by the optimum sensitivity analysis which relates the design variables at each level to the design variables at a higher level.

At the present stage of development, applications beyond the straightforward sensitivity evaluation and extrapolation (see the numerical examples) require further numerical experimentation to assess their usefulness.

Numerical Examples

This section contains an abridged review of the numerical examples provided in Ref. 6 which demonstrate the sensitivity analysis in the context of structural optimization, verify the correctness of the analysis, and provide a measure of usefulness of sensitivity derivatives in estimating the effect of problem parameters on the optimal objective function and design variables. The examples include a three-bar truss for which closed-form solutions are obtained, a ten-bar truss that requires use of a finite-element analysis, and a thin-walled beam characterized by strongly nonlinear constraints of local buckling.

Verification of the Sensitivity Equations

A three-bar truss example that has been referred to in many publications (e.g., Ref. 11) is used herein to verify Eqs. (4) and (15). Numerical data for the truss shown in Fig. 1a are as follows: $L=25.4$ cm (10 in.), $L'=2L$ regardless of the value of angle α, $P=88{,}995$ N (20 kip), the material is an Al alloy, $\sigma_a=13.8$ kN/cm^2 (20 ksi), and $\rho=2.77$ g/cm^3 (0.1 lb/in.^{3}).

Design variables are A_1 and A_2, with $A_3=A_1$. The objective function is material weight $F=\rho L(2\sqrt{2}A_1+A_2)$ and the constraints are on stress in rods 1, 2, and 3 for two loading cases, P_1 and P_2, with $P_1=P_2=P$ and the orientation of the forces P_1 and P_2 not affected by changes of the angle α.

The constraint functions are $g_{uv}=|\sigma_{uv}|/\sigma_a-1<0$ where $\sigma_{11}=P_y\sin\alpha/(L'k)+P_x/(2A_1\cos\alpha)$, $\sigma_{21}=P_y/(Lk)$, $\sigma_{31}=P_y\sin\alpha/(L'k)-P_x/(2A_1\cos\alpha)$, $k=A_1 2\sin^2\alpha/L'+A_2/L$, and $P_x=-(\sqrt{2}/2)P$, $P_y=(\sqrt{2}/2)P$. Subscripts u and v define rod number and load case, respectively. Due to the structure symmetry, only the constraints for stresses shown above need to be included. For the optimum located[11] at a tangent point of a contour of the objective function and the function $g_{11}=0$, one can use the above equations to derive closed-form expressions for optimum A_1 and A_2. For the given data and the particular case of $\alpha=45$ deg, these expressions are $A_1=(P/\sigma_a)(3+\sqrt{3})/6$, $A_2=(P/\sigma_a)1/\sqrt{6}$.

For verification of Eqs. (4) and (15), numerical values of the optimum sensitivity derivatives with respect to P and α were obtained in three ways: from direct differentiation of analytical expressions for A_1 and A_2, from Eqs. (4) and (15). The three sets of derivatives are shown in Table 1 to be in agreement.

Optimization with Conflicting Objectives

For a multiobjective optimization application, the three-bar truss used in the foregoing is modified so that the rods corresponding to cross-section A_1 are assumed to made of steel, $\sigma_a=25.3$ kN/cm^2 (36,667 psi), $\rho=7.81$ g/cm^3 (0.282 lb/in.^{3}), and unit weight cost $c=1.10\times10^{-2}$ \$/g (5\$/lb); the center rod (A_2) is assumed to made of titanium of $\sigma_a=27.6$ kN/cm^2 (40,000 psi), $\rho=4.43$ g/cm^3 (0.160 lb/in.^{3}), and $c=2.43\times10^{-2}$ \$/g (11\$/lb). Minimization of the mass (F_1) and cost (F_2) objectives is considered with the realization that the two objectives will conflict because titanium is lighter but also more costly than steel.

First, an optimization is carried out with a composite objective function $F=a_1F_1+a_2\beta F_2$ from the initial point

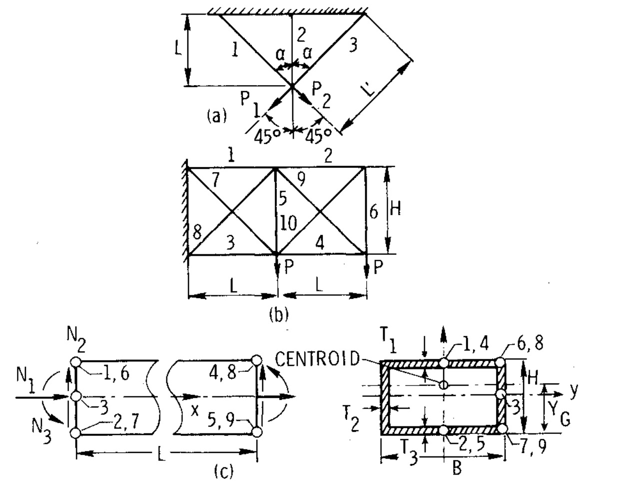

Fig. 1 Numerical examples: a) three-bar truss, b) ten-bar truss, c) box beam.

Table 1 Optima and sensitivity derivatives for the three-bar truss (Fig. 1a)

Variables	Optimal value cm^2	Optimal value in.^{2}	Key[a]	Parameter p: cm^2/N, $\times10^{-5}$	Parameter p: in.^{2}/lb, $\times10^{-5}$	Parameter α: cm^2/rad	Parameter α: in.^{2}/rad
A_1	5.088	0.7887	1	5.713	3.940	−5.700	−0.8835
			2	5.713	3.940	−5.709	−0.8849
			3	5.713	3.940		
A_2	2.641	0.4093	1	2.971	2.049	6.295	0.9757
			2	2.971	2.049	6.318	0.9793
			3	2.974	2.051		
Objective	g	lb		g/N $\times10^{-2}$	lb/lb $\times10^{-4}$	g/rad $\times10^{-2}$	lb/rad
F	1.20×10^3	2.64	1	1.345	1.319	−6.908	−1.523
			2	1.345	1.319	−6.913	−1.524
			3	1.345	1.319		

[a]Key: 1 - from Eq. (4), 2 - from direct differentiation, 3 - from Eq. (15).

$A_1 = A_2 = 1.$, and weighting factors $a_1 = a_2 = 1$ that reflect an approximately even "importance" subjectively assigned to the two objectives (the constant β equalizes the numerical magnitudes of the two terms in F at the initial point). The results of the optimization are given in Table 2, row 1. The reader may find it interesting to compare these results with the optimum solution point given for a single-material truss in Table 1. Next, sensitivity derivatives with respect to weighting ("importance") factors are obtained (Table 2, rows 2 and 3), treating the factors as parameters of the problem, to determine the trends that would be followed should the relative importance of the two objectives change. These trends are extrapolated to the extremes of mass only and cost only in rows 4 and 5 and are verified by the results of full optimizations for each objective separately, starting from the initial point. The objective function extrapolation results shown are obtained by Eq. (18c) which, in this case, provided better accuracy than Eq. (18a). Comparison of rows 4 and 5 with row 1 shows that:

1) Single-objective optimization yields an objective value reduced relative to the value of the same objective obtained from multiobjective optimization (an expected result). In this case, the reductions are small, but significant, 2.52% for the mass and 1.95% for the cost.

2) The extrapolation underpredicts the mass reduction by about 9% and overpredicts the cost reduction by about 36%.

The verification shows the tendency for good prediction of the larger of the design variables and even better prediction for the objective function. Remarkably, the disappearance of the expensive titanium rod in the cost-only optimization is predicted very well by the extrapolation. It appears that the approach has the potential of providing useful estimates of the results of many single-objective optimizations by extrapolating from the results obtained by executing only one optimization with a composite objective function.

Extrapolation of the Optimal Solution

Examples of extrapolation defined by Eqs. (18) are given for a ten-bar truss and a thin-walled beam.

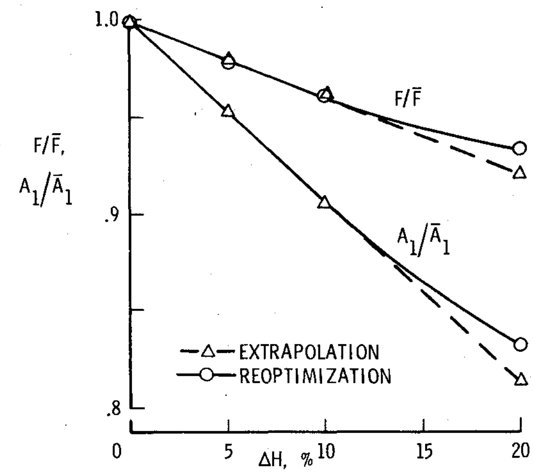

Fig. 2 Ten-bar truss: objective function and one cross section as functions of parameter H.

Table 2 Use of sensitivity and extrapolation for the three-bar truss optimization for mass and cost

	Objective function[a] g	Objective function lb	A_1 cm²	A_1 in.²	A_2 cm²	A_2 in.²
			Optimum for $a_1 = a_2 = 1$			
1	$F_1 = 1.8824 \times 10^3$	4.150	3.168	0.4910	0.955	0.1480
	βF_2[b] $= 1.6766 \times 10^3$	3.6963				
			Sensitivity derivatives, parameter a_1			
2	1.8789×10^3	4.1422	-2.3100×10^{-1}	-3.5805×10^{-2}	7.3361×10^{-1}	1.1371×10^{-1}
			Sensitivity derivatives, parameter a_2			
3	1.6788×10^3	3.7011	2.1288×10^{-1}	3.2997×10^{-2}	-6.7606×10^{-1}	-1.0479×10^{-1}
			Extrapolation and reoptimization for mass only $a_1 = 1$, $a_2 = 0$			
4	1.8392×10^3[c]	4.0547	2.953	0.4577	1.630	0.2526
	1.8351×10^3[d]	4.0457	2.779	0.4308	2.459	0.3812
			Extrapolation and reoptimization for cost only $a_1 = 0$, $a_2 = 1$			
5	1.6320×10^3[c]	3.5980	3.397	0.5265	2.201×10^{-1}	3.411×10^{-2}
	1.6438×10^3[d]	3.6240	3.519	0.5455	6.452×10^{-5}	1.0×10^{-5}

[a] The objective function is $F = a_1 F_1 + a_2 \beta F_2$, F_1 being the mass and F_2 the cost; $\beta = 75.568$ g/$ (0.16660 lb/$). [b] The quantity βF_2, instead of F_2 alone, is referred to consistently in the table. [c] Extrapolation using Eq. (18c). [d] Reoptimization.

Table 3 Optimum solution and its sensitivity derivatives with respect to parameter H for the ten-bar truss (Fig. 1b)

Variable	Optimal value cm²	Optimal value in.²	Sensitivity derivatives cm²/cm	Sensitivity derivatives in.²/in.
1	51.359	7.9606	-5.2507×10^{-2}	-2.0672×10^{-2}
2	0.645	0.1	0.0	0.0
3	52.049	8.0676	-5.1041×10^{-2}	-2.0095×10^{-2}
4	25.497	3.9521	-2.5073×10^{-2}	-9.8714×10^{-3}
5	0.645	0.1	0.0	0.0
6	0.645	0.1	0.0	0.0
7	37.100	5.7505	-1.8192×10^{-2}	-7.1621×10^{-3}
8	36.041	5.5863	-1.9170×10^{-2}	-7.5473×10^{-3}
9	35.935	5.5699	-1.7734×10^{-2}	-6.9817×10^{-3}
10	0.645	0.1	0.0	0.0
Objective	g	lb	g/cm	lb/in.
W	7.239×10^5	1.596×10^3	-3.0444×10^2	-1.70476

Ten-Bar Truss Example

The truss shown in Fig. 1b is another standard (e.g., Ref. 13) test example, chosen herein to illustrate a case too large to be handled by a closed-form solution and to show the usefulness of the sensitivity derivatives in extrapolation. A stiffness-based, finite-element method, augmented with an analytical technique (e.g., Ref. 8, 12, or 13) to generate first and second derivatives of displacements and stresses with respect to the design variables, was used as the analysis program. This program was coupled with a general-purpose optimization program[14] to obtain an optimum solution. The optimum sensitivity algorithm based on Eq. (4) was implemented as a general-purpose program and executed as a postprocessor to the optimization procedure.

The objective function was the material weight, and the 10 cross-sectional areas (numbered in Fig. 1b) were design variables. Constraints were imposed as minimum limits on the variable values and as limits on the stresses. The following detailed data were used: $L=914.4$ cm (360 in.), with $H=L$ initially, one loading case is assumed with $P=444{,}974$ N (10^5 lbf), and the material is an Al alloy of $E=6.9$ MN/cm^2 (10^7 psi), $\sigma_a=17.2$ kN/cm^2 (25 ksi), and $\rho=2.77$ g/cm^3 (0.1 lbf/in.^{3}), $A_{\min}=0.645$ cm^2 (0.1 in.^{2}).

Sensitivity derivatives with respect to the truss depth H are collected in Table 3. The nonzero derivatives are negative, because the increase of H decreases the forces in the horizontal and diagonal rods. Considering variations of H, Fig. 2 presents comparison of reoptimization results for increments of H vs the extrapolation results obtained by Eqs. (18a) and (18b) using sensitivity derivatives from Table 3. The comparison, given in full detail in Ref. 6, is shown herein for the objective function and one typical variable (A_1). The graph indicates that the extrapolation practically coincides with the reoptimization for up to 10% change of H. For a 20% change of H, extrapolation overestimates the decrements of weight and the typical variable by about 18 and 12%, respectively.

For this structure, it would be difficult to predict by physical insight alone how the objective function would change with the increase of H, because the decreases of the horizontal rod cross sections are counteracted by the increases of lengths in the diagonal and vertical rods.

Thin-Walled Box Beam Example

The beam shown in Fig. 1c provides an example with a high degree of nonlinearity because of the constraints which include local buckling and equality constraints imposed on the cross-sectional area and moment of inertia. In this problem the values of variables B, H, T_1, T_2, and T_3 are sought that make A and I equal to prescribed values A_p and I_p while minimizing the objective function $F=\Omega$, where

$$\Omega=\sum_j(\langle g_j\rangle)^2;\ \langle g_j\rangle=\begin{cases} g_j, \text{ if } g_j>0 \\ 0., \text{ if } g_j\leq 0\end{cases} \qquad (22)$$

and the constraints g_j are stress and local buckling constraints. The quantity Ω provides a single measure of unsatisfaction of the constraints and, because of the power factor in Eq. (22), it is continuous up to first derivatives with respect to the design variables, if the constraints are continuous. This continuity is required for the optimizer[14] used in the study. Thus, the optimization problem is

$$\min_{X_i}\ \Omega, \text{ subject to } A=A_p,\ I=I_p, \text{ and } g_j\leq 0 \qquad (23)$$

where inequality constraints g_j are the minimum gage and other geometrical constraints. This somewhat unusual formulation of a structural optimization problem has a meaning of proportioning the detailed dimensions of a cross section to achieve a least violation of the constraints, while conforming to the prescribed cross-sectional stiffness properties in tension (A_p) and bending (I_p), and is motivated by its use in a multilevel decomposition described in the section on Potential Applications. In that decomposition, the A_p and I_p values are to be the system level variable imposed on a beam treated as a subsystem governed by local variables B, H, T_1, T_2, and T_3.

The variables B and T_2 are eliminated from the problem by solving the equality constraint equations $A=A_p$ and $I=I_p$, so that only the variables H, T_1, and T_3 are left as free variables. Similarly, equilibrium equations are used to reduce the number of six end forces to three statically independent end forces. Bending and shear stresses are expressed as functions of the end forces and design variables by means of the usual engineering bending theory formulas and constrained by allowable stress and critical stress of local buckling of the beam walls. The latter are also expressed as functions of the design variables by plate stability formulas from Refs. 15 and 16. The resulting functions, given in detail together with the complete input data in Ref. 6, are algebraically complex; therefore, all the derivative terms in Eq. (4) are computed by finite differences.

The sensitivity derivatives of the independent local variables and the objective function with respect to the bending moment N_3 and the moment of inertia I_p are displayed in Table 4. A typical sample of the extrapolation and reoptimization results for the box beam is plotted in Fig. 3 against the variations of the parameter I_p. The graph shows the capability of the extrapolation to account for practically the entire change of the objective function F and the design variable H (the beam depth) corresponding to a parameter increment of up to 20%. The objective function was extrapolated using Eqs. (18a) and (18c), with the latter giving much better estimates.

The relatively good accuracy of the extrapolation results for the beam problem suggests its potential usefulness in the previously described concept of a multilevel decomposition.

Sensitivity Analysis Accuracy as a Function of the Optimum Solution Degree of Convergence

The relative error of extrapolation obviously depends on the accuracy of the sensitivity derivatives which, in turn, are

Table 4 Optimum solution and derivatives with respect to parameters N_3 (end bending moment) and I_p (moment of inertia corresponding to N_3) for the box beam (Fig. 1c)

Variable (Fig. 1d)	Optimal value, cm (in.)	Derivative $\partial/\partial N_3$, cm/N·cm (in./in.·lb)	Derivative $\partial/\partial I_p$, cm/cm^4 (in./in.^{4})
H	6.806 (2.680)	-1.0396×10^{-6} (-4.624×10^{-6})	4.469×10^{-2} (7.323×10^{-1})
T_1	0.3246 (0.1278)	1.3756×10^{-8} (6.1187×10^{-8})	2.460×10^{-4} (4.031×10^{-3})
T_3	0.5415 (0.2132)	-6.0940×10^{-7} (-2.711×10^{-6})	2.638×10^{-3} (4.323×10^{-2})
B[a]	9.425 (3.711)	—	—
T_2[a]	0.1275 (0.0502)	—	—
		1/N·cm (1/in.·lb)	1/cm^4 (1/in.^{4})
Objective F	0.2735	4.6030×10^{-6} (5.200×10^{-5})	-1.848×10^{-2} (-4.932×10^{-2})

[a] Dependent variables.

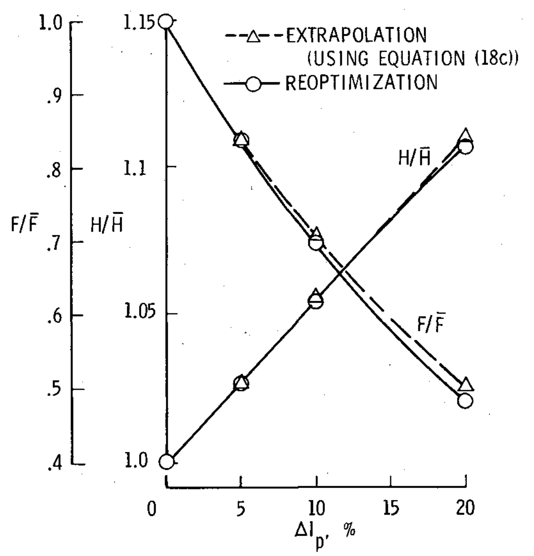

Fig. 3 Box beam: objective function and depth as functions of parameter I_p.

influenced by the degree of convergence of the optimum solution. Although, theoretically, solution convergence is rigorously prescribed by the Kuhn-Tucker conditions, in practice the optimization procedures usually terminate by less rigorous, "practical," criteria. For example, the optimizer[14] used in this study may stop when all of the constraints are satisfied and when the last n iterations (n was assumed to be 5) produced relative differences between two consecutive objective functions smaller than a tolerance C. To be able to use the optimum sensitivity analysis with confidence, one needs to know how strongly the sensitivity derivatives depend on the convergence controls represented by, for example, the tolerance C.

Results that shed some light on that dependence have been given in Ref. 6 for the ten-bar truss and the thin-walled box-beam examples. They showed the influence of C to be rather weak, especially for the objective function derivatives. For example, halving tolerance C in the truss optimization from 0.006 to 0.003 results in only a 0.1% change of the minimum mass derivative with respect to parameter H. In the box-beam case, the objective function derivative with respect to parameter I_p remains practically the same for a tolerance C ranging from 0.025 to 0.001.

Summary of the Numerical Examples

Results of the examples verify the numerical solutions of the sensitivity equations (4) and (15) by analytical solutions for the simple case of a three-bar truss. Additional verification of the sensitivity equations is seen in the agreement of the extrapolation and reoptimization results for small variations of the parameters in the ten-bar truss and the thin-walled beam.

For all cases tested, the extrapolation based on the sensitivity derivatives shows the ability to predict the reoptimization results with a relative error on the order of a few percent for a parameter increment range order of 20%, and the relative error appears to be a rather weak function of the degree of convergence of the optimum solution within practical convergence limits. For smaller increments, below 10%, the relative error of extrapolation is small enough for the accuracy needed in most engineering calculations. This is apparently due to the fact (apparent in Figs. 2 and 3) that, although the relative variations of the optimum objective and variables may be of the same order as the relative increments of the parameters that cause them, they are, at least for the cases tested, nearly linear functions of these parameters despite the nonlinearity of the optimization problems themselves.

The extrapolation predictions are generally better for numerically larger, more significant variables, and better for the objective function than for the variables. In most, but not all, cases tested, the extrapolation via the function recomputation using the extrapolated variables [Eq. (18c)] yielded an accuracy better than extrapolation via the objective function derivatives [Eq. (18a)].

Particularly interesting are the extrapolation results for the shape parameter for the ten-bar truss since they suggest a potential application for decoupling the overall shape variables from the cross-sectional dimension variables in structural optimization.

Still another potential use of the sensitivity analysis is in extrapolation of a single optimization with a composite-objective function to the extremes corresponding to single-objective optimizations, as illustrated by the three-bar truss case with the conflicting objectives of cost and weight.

In the foregoing examples, the extrapolation accuracy benefited from the lack of slightly satisfied constraints with their boundaries near the optimum solution. If such constraints existed, their boundaries could have been penetrated in the process of extrapolation, thus introducing a discontinuity associated with new constraints being brought into the active constraint set. It is obviously important to be alert for such discontinuities, which can be detected by Eq. (21), when extrapolating from an optimum solution point. These potential discontinuities, as well as the degree of nonlinearity of the constraint and objective functions being a characteristic of a particular problem, make the extrapolation accuracy problem dependent. Therefore, its practical applicability range will have to evolve from further accumulation of experience.

Conclusion

Methods for determining the sensitivity derivatives of a constrained minimum solution to the problem parameters are discussed and the governing equations are derived. The equations directly yield the derivatives of the optimum design variables and of the objective function with respect to the parameters that are constants of the problem. For example, derivatives of optimal cross-sectional dimensions and structural mass of a structure can be obtained with respect to allowable stress, load, overall structural dimensions, etc. The derivatives or, in other words, the sensitivity data are obtained by solving a set of linear algebraical equations, thus eliminating the need to repeat the optimization for incremented values of the parameters. While the primary application of the sensitivity analysis is to determine trends at the solution point, it has potential uses in optimization with multiple objectives, estimating changes in the composition of the redundant and active constraint sets, and a formal decomposition of large optimization problems.

The paper refers to structural optimization for verification of the sensitivity solutions and for examples of some of the many potential applications. The examples for a truss demonstrate determination of sensitivity with respect to load and shape parameters. Results are also presented for a ten-bar truss analyzed by a finite-element method. Included as a test structure is a thin-walled beam that introduces a high degree of nonlinearity through local buckling constraints. The results show that, at least for the example problems, a practically significant extrapolation accuracy may be obtained for a reasonably broad range of parameter changes. The results also show that accuracy does not depend strongly on the degree of convergence of the optimum solution from which the sensitivity derivatives are obtained as long as that optimum is within practical convergence bounds.

An example of optimization of a two-material truss with conflicting objectives of mass and cost points to the usefulness of the sensitivity analysis to predict trends and to extrapolate with respect to the individual objectives from the basis of a single optimization with a composite-objective function.

Finally, while the examples used in the paper are structural optimization problems, the equations for the sensitivity derivatives are entirely general since they are derived from the basic Lagrange multiplier equations or, alternatively, from extremum conditions for a penalty function. Consequently, the optimum sensitivity analysis applies to optimization problems in any discipline and also to multidisciplinary systems.

References

[1]Garvin, W. W., *Introduction to Linear Programming*, McGraw-Hill Book Co., New York, 1960.

[2]Fiacco, A. V. and McCormick, G. P., *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley & Sons, New York, 1968, Sec. 2.4.

[3]Fiacco, A. V., "Sensitivity Analysis for Nonlinear Programming Using Penalty Methods," *Mathematical Programming*, Vol. 10, 1976, pp. 287-311.

[4]Armacost, R. L. and Fiacco, A. V., "Computational Experience in Sensitivity Analysis for Nonlinear Programming," *Mathematical Programming*, Vol. 6, 1974, pp. 301-326.

[5]McCormick, G. P., "Algorithmic and Computational Aspects of the Use of Optimization Methods in Engineering Design," *Computers and Structures*, Vol. 3, No. 5, Sept. 1973, pp. 1241-1250.

[6]Sobieszczanski-Sobieski, J., Barthelemy, J.-F., and Riley, K. M., "Sensitivity of Optimum Solutions to Problem Parameters," NASA TM 83134, May 1981; also, AIAA Paper 81-0548, *Proceedings, AIAA/ASME/ASCE/AHS 22nd Structures, Structural Dynamics and Materials Conference*, Atlanta, 1981, pp. 184-205.

[7]Noor, A. K. and Peters, J. M., "Reduced Basis Technique for Nonlinear Analysis of Structures," AIAA Paper 79-0747, April 1979.

[8]Fox, R. L., *Optimization Methods for Engineering Design*, Addison-Wesley, Reading, Mass., 1971.

[9]Haftka, R. T. and Starnes, J. H. Jr., "Applications of a Quadratic Extended Interior Penalty Function for Structural Optimization," *AIAA Journal*, Vol. 14, June 1976, pp. 718-724.

[10]Haftka, R. T., "Automated Procedure for Design of Wing Structures to Satisfy Strength and Flutter Requirements," NASA TN D-7264, July 1973.

[11]Schmit, L. A. and Pope, G. G., "Introduction and Basic Concepts: Structural Design Applications of Mathematical Programming Techniques," AGARDograph 149, 1971, Chap. 1.

[12]Storaasli, O. O. and Sobieszczanski, J., "On the Accuracy of the Taylor Approximation for Structure Resizing," *AIAA Journal*, Vol. 12, Feb. 1974, pp. 231-233.

[13]Schmit, L. A. Jr. and Farshi, B., "Some Approximation Concepts for Structural Synthesis," *AIAA Journal*, Vol. 12, May 1974, pp. 692-699.

[14]Vanderplaats, G. N., "The Computer Design and Optimization," *Computing in Applied Mechanics*, edited by R. F. Hartung, AMD Vol. 18, ASME, New York, 1976, pp. 25-48.

[15]Timoshenko, S. P. and Gere, J. M., *Theory of Elastic Stability*, McGraw Hill Book Co., New York, 1961.

[16]Bruhn, E. F., *Analysis and Design of Flight Vehicle Structures*, Tri State Offset Co., Cincinnati, Ohio, 1965.