Kozłowski, W., and Mróz, Z. (1969), Optimal design of solid plates, *Int. J. Solids Struct.*, 5, 781-794

Mróz, Z., Kamat, M.P., and Plaut, R.H. (1985), Sensitivity analysis and optimal design of non-linear beams and plates, *J. Struct. Mech.*, 13, 245-266.

Mróz, Z. (1987), Sensitivity analysis and optimal design with account for varying shape and support conditions, Proc. NATO ASI *Computer Aided Optimal design: Structural and Mechanical Systems*, Ed. C.A. Mota Soares, *Springer Verlag*, 407-439.

Rozvany, G.I.N., Olhoff, N., Cheng, K.T., and Taylor, J. (1982), On the solid plate paradox in structural optimization, *J. Struct. Mech.*, 10, 1-32.

Rozvany, G.I.N., Ong, T.G., Szeto, W.T., Sandler, R., Olhoff, N., and Bendsøe, M.P. (1987), Last-weight design of perforated elastic plates, *Int. J. Solids Struct.*, Part I., 23, (4), 521-536, Part II., 23, (4), 537-550.

Samsonov, A.M. (1978), Optimal position of an elastic thin rib on elastic plate, *Mech. Tv. Tela*, 1, 132-138.

DUAL METHODS FOR CONVEX SEPARABLE PROBLEMS

C. FLEURY
Aerospace Laboratory
University of Liege
Rue E. Solvay, 21
B-4000 Liege, Belgium

ABSTRACT. In this Lecture, dual methods for solving constrained optimization problems are presented. These methods proceed by maximizing a dual function which depends only on the Lagrangian multipliers associated with the constraints. The Lagrangian multipliers, also called dual variables, are required to remain non-negative. The dual method approach is specially useful and efficient when dealing with convex, separable problems.

In order to apply dual methods to general problems, a successful strategy consists of using convex approximation schemes. In this "Sequential Convex Programming" (SCP) approach, the primary optimization problem is replaced with a sequence of convex explicit subproblems having a simple algebraic form. Various SCP techniques have recently emerged, that have demonstrated strong potential for efficient solution of structural optimization problems. To illustrate the SCP approach, emphasis will be placed on the "Convex Linearization" (CONLIN) method, because it leads to a relatively simple dual formulation.

Next, attention is focused on convex, separable, quadratic problems for which the dual function can be written explicitly as a quadratic (but not separable) form. Efficient second order algorithms based on update formulas for the inverse Hessian matrix are presented. It is finally shown how general separable problems can be solved as a sequence of quadratic subproblems.

1. Introduction

This Lecture is focused on the use of the dual method approach to solve general constrained problems arising in structural optimization. These nonlinear programming problems can of course be solved iteratively by using primal optimization techniques such as those described in another Lecture. Each iteration begins with a complete analysis of the system behaviour in order to evaluate the objective function and constraint values along with their sensitivities to changes in the design variables (i.e. first derivatives). Most often the analysis capability is based on finite element discretization. A design iteration is concluded by employing the results of these behavioral and sensitivity analyses in a minimization algorithm which searches the n-dimensional design space for a new primal point that decreases the objective function value while remaining feasible (i.e. satisfying the constraints).

The essential difficulty in solving directly structural optimization problems via primal methods lies in the implicit character of the functions that define the problem. In other words, for each new design, these functions can only be evaluated numerically through

G. I. N. Rozvany (ed.), Optimization of Large Structural Systems, Vol. I, 509–530.

a finite element analysis. The iterative nature of the optimization process implies that many structural reanalyses must usually be accomplished before finding an acceptable solution. Those repeated analyses can lead to a prohibitive computational cost when dealing with large scale problems.

On the other hand, in the dual method approach, the constrained primal minimization problem is replaced by the maximization of a quasi-unconstrained dual function depending only on the Lagrangian multipliers associated with the constraints. These multipliers are the dual variables subject to simple non-negativity constraints. The dual method approach is well known and quite respected in the mathematical programming community [1,2]. In the context of structural optimization problems, it was initially introduced in [3], and it subsequently led to a reconciliation of optimality criteria techniques and mathematical programming methods [4]. The efficiency of the dual formulation is due to the fact that maximization is performed in the dual space, whose dimensionality is relatively low and depends on the number of active constraints at each design iteration.

It is essential to point out that duality concepts are best exploited for problems presenting special properties, including convexity and separability. Consequently, in order to effectively apply dual methods to general problems, specially devised "linearization" techniques must be used. A successful strategy consists of using convex approximation schemes. In this approach, that can be named "Sequential Convex Programming" (SCP), the primary optimization problem is replaced with a sequence of explicit approximate subproblems having a simple algebraic form. A very attractive feature of the SCP approach is that, because each subproblem is convex and separable, it can be readily solved by a dual method formulation.

Various SCP techniques have recently emerged, that have demonstrated strong potential for efficient solution of structural optimization problems. To illustrate the SCP approach, emphasis will be placed on the "Convex Linearization" (CONLIN) method, because it leads to a relatively simple dual formulation. In the convex linearization method, the initial problem is transformed into a sequence of explicit subproblems having a quite simple algebraic structure. Furthermore each subproblem is convex and separable. These properties make it attractive to solve the subproblem by using dual algorithms.

Turning to the algorithmic point of view, attention will be focused on convex, separable, quadratic problems for which the dual function can be written explicitly as a quadratic (but not separable) form. Efficient second order algorithms based on update formulas for the inverse Hessian matrix are presented. It is finally shown how general separable problems can be solved as a sequence of quadratic subproblems.

2. Primal and Dual Problems

2.1. THE LAGRANGIAN MULTIPLIER TECHNIQUE (EQUALITY CONSTRAINTS)

The origin of the dual method approach can probably be traced back to the classical Lagrangian multiplier technique for minimization problems involving equality constraints:

$$\text{minimize} \quad f(x)$$
$$\text{subject to} \quad c_j(x) = 0 \qquad j=1,m \tag{1}$$

where x represents a point in the n-dimensional Euclidian space E^n, $f(x)$ is the objective function, and $c_j(x)$ are the constraint functions (which are supposed to be linearly independent). Let us assume that this problem has at least one solution x^*. It is well known that, under rather unrestrictive conditions, there exists a vector $\lambda^* \in E^m$ of Lagrangian multipliers such that (x^*, λ^*) is a stationarity point of the Lagrangian function

$$L(x,\lambda) = f(x) - \Sigma\, \lambda_j\, c_j(x) \tag{2}$$

The stationarity conditions state that:

$$\nabla_x L(x,\lambda) = 0 \tag{3}$$

$$\nabla_\lambda L(x,\lambda) = 0 \tag{4}$$

Therefore a stationary point is a solution of the n + m nonlinear equations:

$$\partial f(x)/\partial x_i - \sum_j \lambda_j\, \partial c_j(x)/\partial x_i = 0 \qquad i=1,n \tag{3'}$$

$$c_j(x) = 0 \qquad j=1,m \tag{4'}$$

with n + m unknowns $x_1,\dots, x_n, \lambda_1,\dots, \lambda_m$.

The Lagrangian multiplier technique consists of solving the system of equations (3,4) in order to obtain a solution for the constrained minimization problem (1). Obviously, this procedure increases the dimensionality of the problem. However, in some special cases, the n equations (3) can provide explicitly the variables x_i in terms of the Lagrangian multipliers λ_j, which then become the only unknowns; they are the solution of the system of nonlinear equations:

$$c_j[x(\lambda)] = 0 \qquad j=1,m$$

As an illustration, let us consider a quadratic problem with linear equality constraints:

$$\text{minimize} \quad \tfrac{1}{2}\, x^T x$$
$$\text{subject to} \quad C^T x = d \tag{5}$$

where C denotes the nxm matrix of the constraint gradients (which are assumed linearly independent). The Lagrangian function

$$L(x,\lambda) = \tfrac{1}{2}\, x^T x - \lambda^T (C^T x - d)$$

being fully explicit, the stationarity conditions lead to a closed-form solution of the problem, as shown below:

$$\nabla_x L \equiv x - C\,\lambda = 0 \quad ===> \quad x = C\lambda \tag{3''}$$

$$\nabla_\lambda L \equiv C^T x - d = 0 \quad ===> \quad C^T x = d \tag{4''}$$

Inserting (3") into (4") yields the optimal values of the Lagrangian multipliers:

$$\lambda^* = (C^TC)^{-1}d$$

from which the optimal values of the design variables can be recovered:

$$x^* = C\,\lambda^* = C(C^TC)^{-1}d$$

To introduce the concept of dual problem, it is useful to continue further on this example, and to observe that the relation $x = C\lambda$ resulting from the stationarity conditions can be interpreted as expressing the design variables x in terms of the Lagrangian multipliers λ. Now, inserting $x = C\lambda$ into the Lagrangian function yield a function which depends only on λ:

$$\ell(\lambda) = \tfrac{1}{2}\,\lambda^TC^TC\lambda - \lambda^T(C^TC\lambda - d)$$

$$= -\tfrac{1}{2}\,\lambda^TC^TC\lambda + \lambda^Td$$

This is precisely the dual function. As it will be shown later, the optimal values of the Lagrangian multipliers can be obtained by maximizing this function.

2.2. DUALITY FOR CONVEX PROBLEMS (INEQUALITY CONSTRAINTS)

In the more general case of inequality constraints, the Lagrangian multiplier technique can be generalized provided that adequate convexity conditions are stated. Although duality concepts can be used (with caution!) under weaker convexity properties, we shall simply assume that the "primal problem" to be solved,

$$\text{minimize} \quad f(x)$$
$$\text{subject to} \quad c_j(x) \geq 0 \qquad j=1,m \tag{6}$$

is convex, that is, the objective function as well as the constraint functions are all convex. Then the well known Kuhn-Tucker conditions are necessary and sufficient for global optimality. These conditions can be thought of as a generalization to inequality constraints of the stationarity conditions (3,4). They state that x^* is the global minimum of the convex programming problem (6) if and only if there exist scalars $\lambda_1^*, \ldots, \lambda_m^*$ such that:

$$\partial f(x^*)/\partial x_i - \sum_j \lambda_j^*\, \partial c_j(x^*)/\partial x_i = 0 \qquad i=1,n$$

$$c_j(x^*) \geq 0 \qquad j=1,m$$

$$\lambda_j^*\, c_j(x^*) = 0 \qquad j=1,m$$

$$\lambda_j^* \geq 0 \qquad j=1,m$$

The non-negative scalars λ_j are called generalized Lagrangian multipliers or dual variables.

The first of these conditions implies that x^* is the unconstrained global minimum of the convex Lagrangian function $L(x,\lambda^*)$. Hence it is readily shown that there is a unique correspondence between the primal variables x and the dual variables λ, through the following unconstrained minimization problem, the so-called "Lagrangian problem":

$$\underset{x}{\text{minimize}}\ L(x,\lambda)$$

for which the optimality conditions are:

$$\partial L(x,\lambda)/\partial x_i = 0 \qquad i=1,n$$

Note that the minimization in the above Lagrangian problem is taken over x for any fixed λ. This permits defining the dual function, which depends solely on the dual variables λ_j:

$$\ell(\lambda) = \min_x L(x,\lambda)$$

Alternatively, if $x(\lambda)$ denotes the primal point minimizing the Lagrangian function for a given dual point λ, the dual function can be written:

$$\ell(\lambda) = L[x(\lambda),\lambda] = f[x(\lambda)] - \sum \lambda_j\, c_j[x(\lambda)] \tag{7}$$

It can be shown that the dual function is concave. In addition, for any feasible primal point x, and any feasible dual point λ, then:

$$\ell(\lambda) \leq f(x)$$

The "dual problem" consists of maximizing the dual function subject only to non-negativity constraints on the dual variables:

$$\text{maximize} \quad \ell(\lambda)$$

$$\text{subject to} \quad \lambda_j \geq 0$$

It is useful to repeat here that we assume the primal problem to be convex. Under this condition, its dual is convex too and the respective solutions of both problems satisfy the same optimality conditions. As a consequence, the two problems are equivalent, in the sense that the solution to one provides a solution to the other. Furthermore their optimal values are equal:

$$\ell(\lambda^*) = f(x^*)$$

Probably the most interesting property of the dual function is that its first partial derivatives are easily available. They are given by the negative of the primal constraints:

$$\partial \ell/\partial \lambda_j = - c_j[x(\lambda)] \qquad (8)$$

The optimality conditions for the quasi-unconstrained dual problem can be written:

$$\partial \ell/\partial \lambda_j \equiv c_j[x(\lambda)] = 0 \qquad \text{if } \lambda_j > 0$$

$$\partial \ell/\partial \lambda_j \equiv -c_j[x(\lambda)] < 0 \qquad \text{if } \lambda_j = 0$$

These conditions indicate that the maximum of the dual function will be attained when the constraints are satisfied as equalities for positive dual variables, and as inequalities for zero dual variables. Going a bit further, let us assume that a steepest ascent algorithm is employed to maximize the dual function, i.e.,

$$\lambda = \lambda + \alpha \nabla \ell,$$

or

$$\lambda_j = \lambda_j - \alpha c_j$$

where α denotes a step size. Thus a dual variable λ_j increases if the corresponding constraint c_j is violated, while it decreases (possibly reaching zero) if c_j is positive. From these considerations results an intuitive interpretation of the dual method approach: this approach attempts to satisfy the inequality constraints by adjusting the values of the dual variables.

Because non-negativity constraints are very simple to take into account, classical unconstrained maximization techniques can readily be adapted to solve the quasi-unconstrained dual problem. In particular, the conjugate gradient method is well suited: evaluating the dual function (7) demands the computation of the constraint values $c_j[x(\lambda)]$, so that the gradient (8) is directly obtained without additional computations.

2.3. DUALITY FOR SEPARABLE PROBLEMS

The main difficulty in the dual approach is that, at each point in the dual space, it is necessary to find the x that minimizes the Lagrangian function $L(x,\lambda)$ for fixed λ. This minimization problem in the primal space must be repeated many times, and this might well lead to a prohibitive computational cost. However, for certain problems, for example separable programming problems, this is not very cumbersome. Therefore, in addition to convexity, separability is an essential property for the dual formulation to be efficient.

A function is said to be separable if it can be written

$$c(x) = \Sigma\, c_i(x_i)$$

where each function $c_i(x_i)$ depends only on the single variable x_i. Separable functions benefit from some computationally important properties. In particular, the Hessian matrix of such a function is diagonal. Several examples of separable functions will be encountered later in this Lecture, as well as in other Lectures. Now, the constrained optimization problem (6) is a separable programming problem if each function $\{f(x), c_j(x)\}$ is itself separable. This implies that the Lagrangian function is separable too. As a result, the n-dimensional Lagrangian problem can be broken up into n one-dimensional minimization problems, and the dual function can be written as follows:

$$\ell(\lambda) = \sum_i \min_{x_i} [f_i(x_i) - \sum_j \lambda_j c_{ij}(x_i)]$$

In many cases, the single-variable minimization problem appearing above has a simple algebraic structure and it can be solved in closed form, yielding thus an explicit dual function.

As an illustration, let us reconsider the quadratic problem (5), but now with linear inequality constraints:

<u>Primal Problem</u>

minimize $\quad \frac{1}{2} x^T x$

subject to $\quad C^T x \geq d$

<u>Lagrangian Problem</u>

minimize $\quad L(x,\lambda) = \frac{1}{2} x^T x - \lambda^T (C^T x - d)$

To obtain the solution of this convex minimization problem, it is sufficient to state that the gradient of the Lagrangian function must vanish:

$$\nabla L = 0 \implies x(\lambda) = C\lambda$$

The dual function is obtained by inserting these fully explicit primal-dual relationships into the expression of the Lagrangian function:

$$\ell(\lambda) = \frac{1}{2} x^T(\lambda)\, x(\lambda) - \lambda^T [C^T x(\lambda) - d]$$

$$= \frac{1}{2} \lambda^T C^T C \lambda - \lambda^T (C^T C \lambda - d)$$

$$= - \frac{1}{2} \lambda^T C^T C \lambda + \lambda^T d$$

The gradient of the dual function is given by:

$$\nabla \ell(\lambda) = - C^T C \lambda + d$$

$$= d - C^T x(\lambda)$$

that is, by the negative of the primal constraints, in agreement with the general theory of duality.

From this latter expression, it follows that the dual Hessian matrix is constant:

$$\nabla^2 \ell(\lambda) = - C^T C$$

The dual function is therefore quadratic. It has the explicit form:

$$\ell(\lambda) = -\tfrac{1}{2}\lambda^T A \lambda + \lambda^T d$$

where $A = C^T C$ denotes the negative of the Hessian matrix.

In the case of equality constraints, the dual variables are unrestricted in sign. Therefore the maximum of the dual function can simply be obtained by stating that its gradient must vanish:

$$\nabla\ell(\lambda) \equiv -A\lambda + d = 0$$

leading to

$$\lambda^* = A^{-1} d$$

From this dual solution, the primal optimum is:

$$x^* = C\,\lambda^*$$

These dual and primal solutions are of course the same as those generated by the stationarity conditions of the Lagrangian function in Section 2.1.

Now, in the case of linear inequality constraints, the dual variables are required to remain non-negative, and the dual problem becomes:

<u>Dual Problem</u>

$$\text{maximize} \quad \ell(\lambda) = -\tfrac{1}{2}\lambda^T A \lambda + \lambda^T d$$

$$\text{subject to} \quad \lambda_j \geq 0$$

After solving this maximization problem in the dual space, the primal optimum can still be obtained as $x^* = C\,\lambda^*$

Because of the non-negativity conditions that the dual variables must fulfil, a direct solution from the stationarity conditions becomes difficult, if not impossible. For now almost two decades, the so-called "Optimality Criteria" technique has attempted to achieve such an easy direct solution, so far without any significant success. Maximizing the dual function, on the other hand, provides a quite natural and efficient alternative. An algorithm that is specially devised for that purpose is discussed in Section 4.1.

2.4. EXAMPLE: QUADRATIC PROBLEM

To illustrate the duality concepts explained above, we consider the following problem:

$$\text{minimize} \quad \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2$$

$$\text{subject to} \quad x_1 + x_2 \geq 4$$

$$x_1 - x_2 \geq -4$$

The Lagrangian function for this explicit problem has the form:

$$L(x,\lambda_1,\lambda_2) = \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2 - \lambda_1(x_1+x_2-4) - \lambda_2(x_1-x_2+4)$$

It is easily verified that the Lagrangian problem leads here to linear primal-dual relationships:

$$x_1(\lambda_1,\lambda_2) = \lambda_1 + \lambda_2$$

$$x_2(\lambda_1,\lambda_2) = \lambda_1 - \lambda_2$$

The dual function can now be obtained by substituting $x_1(\lambda_1,\lambda_2)$ and $x_2(\lambda_1,\lambda_2)$ in the definition of the Lagrangian function. This yields the following explicit dual problem:

$$\text{maximize} \quad \ell(\lambda_1,\lambda_2) = -(\lambda_1 - 2)^2 - (\lambda_2 + 2)^2 + 8$$

$$\text{subject to} \quad \lambda_1 \geq 0$$

$$\lambda_2 \geq 0$$

As predicted by the duality theory, the first derivatives of the dual function are directly related to the primal constraints:

$$\partial\ell/\partial\lambda_1 = -2\lambda_1 + 4 \equiv 4 - x_1 - x_2$$

$$\partial\ell/\partial\lambda_2 = -2\lambda_2 - 4 \equiv -4 - x_1 + x_2$$

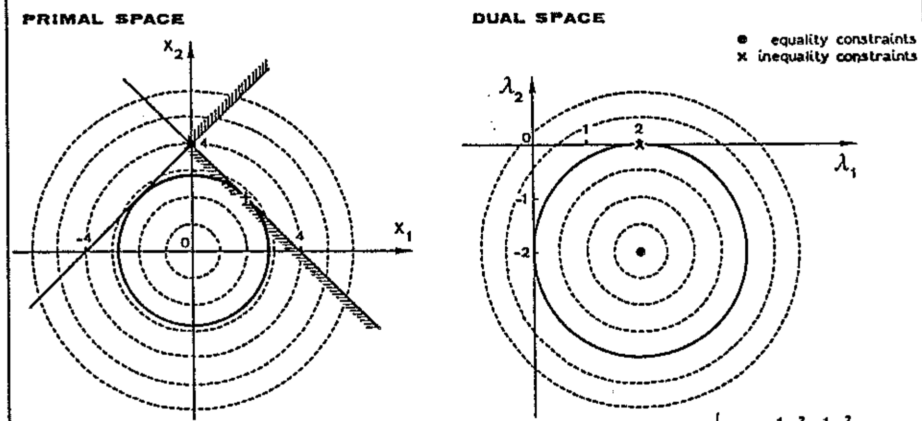

FIG. 1 EXAMPLE OF SEPARABLE QUADRATIC PROBLEM — min $\frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$ s.t. $x_1+x_2 \geq 4$; $x_1-x_2 \geq -4$

A graphical representation of both the primal and dual spaces for this problem is shown in Fig. 1. It can be immediately concluded that the dual optimum lies at $\lambda^* = [2,0]^T$ (with

$l(\lambda^*) = 4$), corresponding to the primal solution $x^* = [2,2]^T$ (with $f(x^*)=4$). Note that the maximum value of the dual function coincides with the minimum value of the primal function. Consider now the same problem, but with equality instead of inequality constraints. The dual formulation derived above remains valid, except that the nonnegativity constraints on the dual variables must be removed. From Fig. 1, it is apparent that the unconstrained maximum of the dual function lies at $\lambda^* = [2,-2]^T$, corresponding to the primal solution $x^* = [0,4]^T$. It is worth noticing that the optimal values of the primal and dual functions are again identical ($f(x^*) = l(\lambda^*) = 8$).

The reader is encouraged to work deeply through the details of this example in order to better understand the dual approach. It is especially recommanded to resort to the matrix formulation previously presented for separable quadratic problems. In this specific case, the primal problem is described by the matrix

$$C = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

while the dual problem is defined by the matrix

$$A = C^T C = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

3. Sequential Convex Programming

Attention is now turned toward the use of the dual method approach to solve general structural optimization problems which do not inherently exhibit an explicit separable form. Such problems can be written in the following form:

$$\text{minimize} \quad c_0(x) \tag{9}$$

$$\text{subject to} \quad c_j(x) \leq 0 \qquad j=1,m \tag{10}$$

$$\underline{x}_i \leq x_i \leq \overline{x}_i \qquad i=1,n \tag{11}$$

where each function $c_j(x)$ depends implicitly on the design variables. In other words, for each new design, these functions can only be evaluated numerically, for example through a finite element analysis.

For reasons that will become clear later, the objective function is now denoted $c_0(x)$ instead of $f(x)$. Also, we have intentionally changed the sense of the inequality sign in the main constraints (10), often called behavior constraints. This is because they usually impose upper limits on structural response quantities such as stresses and displacements. The reader is invited to check the incidence of this change on the dual formulation presented below (expression of the Lagrangian function, of the dual gradient, etc...) It should be noted that the so-called side constraints (11) constitute a particular case of the more general constraints (10). However they are written separately in our optimization problem statement, because the dual approach described below can handle them more efficiently when considered apart from the behavior constraints.

This Section is concerned with various methods based on explicit approximation schemes, that have demonstrated a strong potential for efficient solution of design optimization problems. In these methods, the key idea is to replace the implicit problem (9-11) with a sequence of convex explicit subproblems having a simple algebraic form. Most of these approximation strategies make the assumption that the objective function and constraint functions are separable in terms of the design variables. Such an assumption is not new. During the last two decades, it has been extensively used in several research efforts dealing with structural optimization problems [4-7].

For example, it has now been well accepted that the best approaches to optimal sizing problems are those which make use of constraint linearization with respect to the reciprocal design variables. As demonstrated in Ref. [4], linearizing the behavior constraints in terms of the reciprocal variables was at the basis of both the optimality criteria approach to structural optimization [5], and the mathematical programming approach based on approximation concepts [6,7]. There is an intuitive explanation for the success of this idea, in that stresses and displacements are exact linear functions of the reciprocal sizing variables in the case of a statically determinate structure. Although there is no reason to believe that the use of reciprocal variables constitutes a universal way of solving structural optimization problems, surprisingly, this very special linearization technique has been highly successful in many cases. For shape optimal design problems, for example, there is no clear physical guideline for the selection of intermediate linearization variables. Nevertheless, the choice of reciprocal variables continues to have a beneficial effect on the convergence properties of the optimization process [8,9].

The key idea in the "Convex Linearization" method (CONLIN) [10-12] is to perform the linearization process with respect to mixed variables, either direct or reciprocal, independently for each function involved in the optimization problem. At each successive iteration point, the CONLIN method only requires evaluation of the objective and constraint functions and their first derivatives with respect to the design variables. The optimizer will then select by itself an appropriate approximation scheme on the basis of the signs of the derivatives.

The CONLIN method proceeds by linearizing each function defining the optimum design problem with respect to a properly selected mix of direct and reciprocal variables, so that a convex and separable subproblem is generated. The selection of the "intermediate" linearization variables is made on the basis of the signs of the first partial derivatives. It is easily proven that, considering any differentiable function $c(x)$, the following linearization scheme yields a convex approximation (hence the appellation "convex linearization" suggested in Ref. [10]):

$$\tilde{c}(x) = c(x^0) + \Sigma_+ \, c_i^0(x_i - x_i^0) - \Sigma_- \, (x_i^0)^2 \, c_i^0(1/x_i - 1/x_i^0) \tag{12}$$

where c_i denote the first derivatives of $c(x)$ with respect to the design variables, evaluated at the current design point x^0:

$$c_i^0 = \partial c(x^0)/\partial x_i$$

The symbol Σ_+ (Σ_-) means "summation over the terms for which c_i^0 is positive (negative)". One of the most interesting feature of the convex linearization scheme is that it also leads to the most conservative approximation amongst all possible combinations of mixed direct/reciprocal variables. This property was initially demonstrated in [13], where conservative approximation was employed to handle difficult buckling constraints.

The CONLIN algorithm applies this convex linearization scheme to the objective function and to all the constraint functions defining the optimization problem. It is convenient to normalize the design variables so that they become equal to unity at the current point x^0 where the problem is linearized:

$$x_i' = x_i/x_i^0 \implies c_i' = c_i^0 \, x_i^0$$

The factor $(x_i^0)^2$ disappears from (12), which becomes:

$$\tilde{c}(x') = c(x^0) + \Sigma_+ \, c_i'(x_i'-1) - \Sigma_- \, c_i'(1/x_i'-1) \qquad (12')$$

Applying the convex linearization technique to each function $c_j(x)$, and dropping the superscript $'$, the following explicit subproblem is generated:

$$\begin{aligned} &\text{minimize} && \Sigma_+ \, c_{i0} \, x_i - \Sigma_- \, c_{i0}/x_i - d_0 \\ &\text{subject to} && \Sigma_+ \, c_{ij} \, x_i - \Sigma_- \, c_{ij}/x_i \le d_j \qquad j=1, m \\ & && \underline{x}_i \le x_i \le \bar{x}_i \end{aligned} \qquad (13)$$

where the c_{ij} 's denote the normalized first derivatives of the objective and constraint functions evaluated at the current point x^0. Note that the constants d_j contain the zero order contributions in the Taylor series expansion:

$$d_j = \sum_i | c_{ij} | \, x_i^0 - c_j(x^0) \qquad j=0, m$$

Applying the dual method outlined in Section 2, the solution of the primal problem (13) can be obtained by the following "Max-Min" two-phase procedure:

$$\text{maximize} \quad \ell(\lambda) = \min_{\underline{x}_i \le x_i \le \bar{x}_i} L(x,\lambda) \qquad (14)$$

The dual function $\ell(\lambda)$ to be maximized results from minimizing the Lagrangian function

$$L(x,\lambda) = \sum_{j=0}^{m} \lambda_j \left(\Sigma_+ \, c_{ij} \, x_i - \Sigma_- \, c_{ij}/x_i - d_j \right)$$

over the acceptable primal variables.

The separability of the primal problem implies that the Lagrangian function can be written as the sum of n individual functions $L_i(x_i)$, and therefore, the n-dimensional minimum problem (14) can be split into n single variable minimization problems:

$$\begin{aligned} &\text{minimize} && L_i(x_i) = a_i \, x_i + b_i/x_i \\ &\text{subject to} && \underline{x}_i \le x_i \le \bar{x}_i \end{aligned} \qquad (15)$$

where a_i and b_i depend linearly on the dual variables λ_j:

$$\begin{aligned} a_i &= \Sigma_+ \, c_{ij} \, \lambda_j \ge 0 \\ \text{and} \quad b_i &= - \Sigma_- \, c_{ij} \, \lambda_j \ge 0 \end{aligned} \qquad (16)$$

For any given set of dual variables, a_i and b_i are fixed coefficients in problem (15). These coefficients remain always non-negative in the feasible region of the dual space ($\lambda_j \ge 0$). Therefore the Lagrangian problem (15) has a unique solution, obtained by stating:

$$L_i'(x_i) \equiv a_i - b_i/x_i^2 = 0 \qquad (17)$$

Since the side constraints (11) must be satisfied, it follows that:

$$x_i = (b_i/a_i)^{1/2} \quad \text{if} \quad \underline{x}_i^2 \le b_i/a_i \le \bar{x}_i^2 \qquad (18)$$

$$x_i = \underline{x}_i \quad \text{if} \quad b_i/a_i \le \underline{x}_i^2 \qquad (19)$$

$$x_i = \bar{x}_i \quad \text{if} \quad \bar{x}_i^2 \le b_i/a_i \qquad (20)$$

Remembering that a_i and b_i depend on the dual variables (see Eq. 16), we have obtained fully explicit primal-dual relationships.

The dual problem can now be expressed in closed form:

$$\max \quad \ell(\lambda) = \sum_{j=0}^{m} \lambda_j \left[\Sigma_+ \, c_{ij} \, x_i(\lambda) - \Sigma_- \, c_{ij}/x_i(\lambda) - d_j \right]$$

$$\text{s.t.} \quad \lambda_j \ge 0$$

where the primal variables x_i are known explicitly in terms of the dual variables λ_j via (18-20). In order for the dual function to be bounded, the dual variables λ_j must be linked by a linear equality constraint. Following a common practice we shall simply assume that the Lagrangian multiplier λ_0 associated with the objective function is assigned a unit value.

A fundamental property of the dual function is that its first derivatives are simply given by the primal constraint values:

$$g_j \equiv \partial\ell/\partial\lambda_j = \Sigma_+ \, c_{ij} \, x_i(\lambda) - \Sigma_- \, c_{ij}/x_i(\lambda) - d_j$$

The dual problem can therefore be solved by applying first order algorithms based, for example, on the conjugate gradient method. Because the Hessian matrix of the dual function is readily available, second order methods can also be employed [11].

Example: linear programming problem

Although the convex linearization method does not pretend to replace the SIMPLEX algorithm, it is capable of solving efficiently a linear programming problem by transforming it into a sequence of (nonlinear) convex subproblems. To illustrate the CONLIN concepts previously described, let us consider this linear programming problem:

$$\text{minimize} \quad c_0(x) \equiv x_1 + 4x_2$$

$$\text{subject to} \quad c_1(x) \equiv x_1 - x_2 \leq 0$$

$$c_2(x) \equiv -3x_1 + 2x_2 \leq -1$$

For this problem, the function first derivatives are constant:

$$c_{10} \equiv \partial c_0/\partial x_1 = 1 \qquad c_{20} \equiv \partial c_0/\partial x_2 = 4$$

$$c_{11} \equiv \partial c_1/\partial x_1 = 1 \qquad c_{21} \equiv \partial c_1/\partial x_2 = -1$$

$$c_{12} \equiv \partial c_2/\partial x_1 = -3 \qquad c_{22} \equiv \partial c_2/\partial x_2 = 2$$

By applying the convex linearization scheme at the initial primal point $x^0 = [3,4]^T$ the following CONLIN subproblem is obtained:

$$\text{minimize} \quad x_1 + 4x_2$$

$$\text{subject to} \quad x_1 + 16/x_2 \leq 8$$

$$27/x_1 + 2x_2 \leq 17$$

Note that the linear objective function is not modified, and that only the terms corresponding to negative first derivatives are affected in the two constraint functions. Let us now write the Lagrangian function for this subproblem:

$$L(x,\lambda) = x_1+4x_2 + \lambda_1(x_1+16/x_2-8) + \lambda_2(27/x_1+2x_2-17)$$

Minimizing this function with respect to x_1 and x_2 for fixed λ_1 and λ_2, we get (see Eq. 17):

$$L_1'(x_1) \equiv 1 + \lambda_1 - 27\lambda_2/x_1^2 = 0$$

$$L_2'(x_2) \equiv 4 - 16\lambda_1/x_2^2 + 2\lambda_2 = 0$$

By solving these equations it can be concluded that:

$$x_1(\lambda_1,\lambda_2) = [27\lambda_2/(1+\lambda_1)]^{½}$$

$$x_2(\lambda_1,\lambda_2) = [8\lambda_1/(2+\lambda_2)]^{½}$$

The dual function can now be obtained by substituting $x_1(\lambda_1,\lambda_2)$ and $x_2(\lambda_1,\lambda_2)$ in the definition of the Lagrangian function. This leads to the following explicit dual problem:

$$\text{maximize} \quad \ell(\lambda_1,\lambda_2) = [27\lambda_2/(1+\lambda_1)]^{½} + [8\lambda_1/(2+\lambda_2)]^{½}$$

$$+ \lambda_1 \{[27\lambda_2/(1+\lambda_1)]^{½} + 16[(2+\lambda_2)/8\lambda_1]^{½} - 8\}$$

$$+ \lambda_2 \{27[(1+\lambda_1)/27\lambda_2]^{½} + 2[8\lambda_1/(2+\lambda_2)]^{½} - 17\}$$

$$\text{subject to} \quad \lambda_1 \geq 0$$

$$\lambda_2 \geq 0$$

It is easily verified that the first derivatives of the dual function are:

$$\partial\ell/\partial\lambda_1 = [27\lambda_2/(1+\lambda_1)]^{½} + 16[(2+\lambda_2)/8\lambda_1]^{½} - 8$$

$$\equiv x_1 + 16/x_2 - 8$$

$$\partial\ell/\partial\lambda_2 = 27[(1+\lambda_1)/27\lambda_2]^{½} + 2[8\lambda_1/(2+\lambda_2)]^{½} - 17$$

$$\equiv 27/x_1 + 2x_2 - 17$$

It is important to emphasize that the foregoing primal-dual relationships, as well as the definition of the dual function and its derivatives, become more complicated when side constraints are imposed in the problem statement. For example if the side constraints

$$4 \geq x_1 \geq 0.5$$

$$4 \geq x_2 \geq 0.5$$

are added, then the dual space is partitioned in several regions separated by second order discontinuity planes. As an exercise, the reader is invited to demonstrate that the equations of these four planes are:

$$\lambda_1 - 108\,\lambda_2 + 1 = 0 \qquad 32\,\lambda_1 - \lambda_2 - 2 = 0$$

$$16\,\lambda_1 - 27\,\lambda_2 + 16 = 0 \qquad \lambda_1 - 2\,\lambda_2 - 4 = 0$$

The definition of the dual problem changes when crossing such a discontinuity plane. The dual solution scheme must therefore take these changes into consideration. Clearly the expressions given above for the dual function and its derivatives are only valid in the region of the dual space where the primal variables are free.

After solving the dual problem, new values are obtained for the primal variables. The CONLIN scheme is then applied again at the new primal point, and an updated explicit subproblem follows, having a similar form. This convex subproblem is replaced with its dual and solved. The full process is repeated until convergence is achieved to the primal point $x^* = [1,1]^T$.

4. Second Order Dual Optimizer

This Section will concentrate on Convex Separable (CS) problems, because many general problems of practical interest can be solved as a sequence of CS subproblems via the use of convex approximation concepts. CONLIN, MMA, and the diagonal SQP method discussed in another Lecture are convincing evidence of the success of the convex approximation approach. Any CS problem can itself be broken up into a sequence of quadratic subproblems, by using second order Taylor series either in the primal space or in the dual space (or in both spaces in a primal-dual approach). Therefore the key is to solve efficiently an explicit primal problem of the following form: minimize a separable quadratic function subject to linear constraints as well as bounds on the design variables.

4.1. DUAL METHOD APPROACH FOR SEPARABLE, QUADRATIC PROBLEMS

Let us therefore consider the following separable quadratic problem, with linear inequality constraints and side constraints:

Primal Problem

minimize $\frac{1}{2} \Sigma x_i^2$

subject to $\Sigma c_{ij} x_i \geq d_j$

$\bar{x}_i \geq x_i \geq \underline{x}_i$

The side constraints, which impose lower and upper bounds on the design variables, can of course be viewed as linear inequality constraints. However this would increase dramatically the number of dual variables, and it is much better to treat the side constraints apart from the general linear constraints.

Lagrangian Problem

$$\min_{\bar{x}_i \geq x_i \geq \underline{x}_i} L(x,\lambda) = \frac{1}{2} \sum_i x_i^2 - \sum_j \lambda_j \left(\sum_i c_{ij} x_i - d_j\right)$$

Because of the separability property, this n-dimensional minimization problem can be split into n one-dimensional problems:

minimize $L_i(x_i) = \frac{1}{2} x_i^2 - (\sum_j c_{ij}\lambda_j)\, x_i$

subject to $\bar{x}_i \geq x_i \geq \underline{x}_i$

Setting to zero the first derivatives of $L_i(x_i)$ yields the following primal-dual relationships:

$$x_i = \Sigma c_{ij}\lambda_j$$

if the side constraints are ignored. Note that, in matrix form, this relation can be written $x = C\lambda$, just as in Section 2.

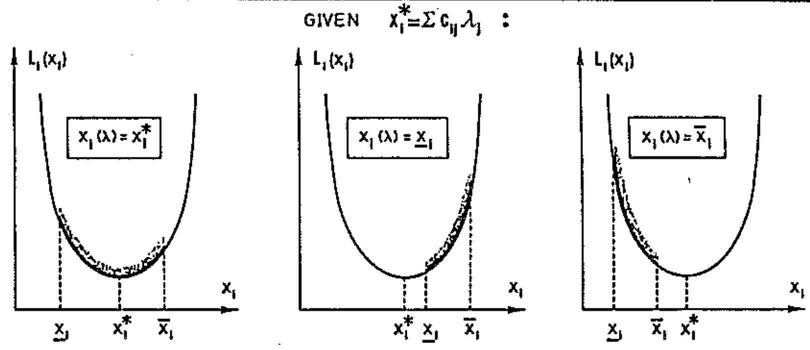

FIG. 2 ILLUSTRATION OF THE LAGRANGIAN PROBLEM

Now, to enforce satisfaction of the side constraints, we need to consider the three possible situations depicted in Fig. 2. It is clear that fully explicit primal-dual relationships $x = x(\lambda)$ are still available. At any point in the dual space, the values of the primal variables can be calculated according to the following steps:

(1) for the given λ, compute $x_i^* = \Sigma c_{ij}\lambda_j$

(2) if $\bar{x}_i > x_i^* > \underline{x}_i$, then $x_i(\lambda) = x_i^*$ (23)

if $x_i^* < \underline{x}_i$, then $x_i(\lambda) = \underline{x}_i$

if $x_i^* > \bar{x}_i$, then $x_i(\lambda) = \bar{x}_i$

The dual function is formed by inserting these primal-dual relationships into the Lagrangian function, which leads to the following statement of the dual problem:

Dual Problem

maximize $\ell(\lambda) = \frac{1}{2} \sum_i x_i^2(\lambda) - \sum_j \lambda_j \left(\sum_i c_{ij}\, x_i(\lambda) - d_j\right)$

subject to $\lambda_j \geq 0$

The first derivatives of the dual function are simply given by the primal constraints, in agreement with the general theory of duality:

$$\partial l/\partial\lambda_j \equiv d_j - \Sigma c_{ij}\, x_i(\lambda)$$

The second derivatives of the dual function, yielding the Hessian matrix (denoted A as in Section 2.2), are easily available too:

$$A_{jk} \equiv \partial^2 l/\partial\lambda_j\partial\lambda_k = \partial/\partial\lambda_k[d_j - \Sigma c_{ij}\, x_i(\lambda)] = - \Sigma c_{ij}\, \partial x_i/\partial\lambda_k$$

But now, instead of the simple expression $A = -C^TC$ derived in Section 2.2, a slightly more complicated situation arises, because of the separate treatment of the side constraints. On one hand, for the free variables, we have:

$$\partial x_i/\partial\lambda_k = c_{ik} \quad \text{if} \quad \bar{x}_i > x_i > \underline{x}_i$$

On the other hand, for the fixed variables, it is obvious that:

$$\partial x_i/\partial\lambda_k = 0 \quad \text{if} \quad x_i = \bar{x}_i \text{ or } x_i = \underline{x}_i$$

Therefore, the Hessian matrix takes on the form:

$$A_{jk} \equiv \partial^2 l/\partial\lambda_j\partial\lambda_k = -\sum_{free} c_{ij}\, c_{ik} \tag{24}$$

where the summation is restricted to the free primal variables.

The second derivatives of the dual function are therefore discontinuous whenever a free primal variable becomes fixed, or conversely. From the primal-dual relationships (23) it is clear that the dual space is partitioned in several regions separated by second order discontinuity planes. These planes are defined by:

$$\Sigma\, c_{ij}\lambda_j = \underline{x}_i \qquad \Sigma\, c_{ij}\lambda_j = \bar{x}_i$$

4.2. UPDATE FORMULAS FOR THE INVERSE HESSIAN

The fundamental difficulty in using Newton type methods for solving the dual problem resides in the inherent discontinuities of the Hessian matrix. Those discontinuities are due to the side constraints imposed on the primal variables. Fortunately the topology of the dual space can be described in an exact mathematical way via the concept of second order discontinuity planes. Based on this concept, a very efficient and reliable algorithm can be devised to solve the dual problem.

The new algorithm works in the dual space, and it is based on the following observation: the Hessian matrix of the dual function is mathematically expressed as a sum of rank one matrices, each matrix corresponding to a free primal variable (i.e. a variable that is not currently fixed to its lower or its upper bound); Eq. (24) can indeed be rewritten in matrix form:

$$A = -\sum_{free} w_i w^T_i$$

where w_i denotes row (i) of the matrix C (i.e. $w_j = c_{ij}$).

It is therefore possible to gradually build up the inverse of the dual Hessian through a series of rank one updates. The algorithm is of course iterative in nature. At each iteration four possible situations can arise:

(1) a free primal variable becomes fixed, and consequently one rank one term must be removed from the dual Hessian;

(2) conversely a fixed primal variable may become free, and it will now contribute to the dual Hessian;

(3) a violated primal linear constraint must be activated, which means that the corresponding zero dual variable will become positive; the dimension of the dual Hessian is then increased by one;

(4) conversely a dual variable may reach the zero value (the corresponding primal constraint becomes inactive); in this case the dimension of the dual Hessian is decreased by one.

In each of these four cases, it is possible to directly compute the inverse dual Hessian via an appropriate rank one update formula (dyadic product). In fact, the dual Hessian matrix itself is not required in the algorithm, and only its inverse is actually evaluated through a series of rank one updates.

Update formula #1 (activated side constraint)

Assume that the primal variable x_i becomes fixed and let w be the corresponding row (i) of the matrix C (i.e. $w_j = c_{ij}$). Then the new Hessian is:

$$\tilde{A} = A + ww^T$$

and its inverse is updated by the formula:

$$\tilde{A}^{-1} = A^{-1} - vv^T/\gamma \tag{25}$$

where

$$v = A^{-1}w$$

$$\gamma = w^Tv - 1$$

Update formula #2 (deactivated side constraint)

Assume now that the primal variable x_i becomes free and let again w be the corresponding row (i) of the matrix C (i.e. $w_j = c_{ij}$). Then the new Hessian is:

$$\tilde{A} = A - ww^T$$

and its inverse is updated by the same formula (25), but with

$$\gamma = w^Tv + 1$$

Update formula #3 (activated linear constraint)

Assume that the primal linear constraint c_k becomes active and let h be the corresponding column (k) of the matrix C (i.e. $h_i = c_{ik}$). Then the new Hessian is:

$$\tilde{A} = \begin{bmatrix} A & w \\ w^T & \alpha \end{bmatrix}$$

with $w_j = -\Sigma c_{ij} h_i$

$\alpha = -\Sigma h_i^2$

Its inverse is updated by the formula:

$$\tilde{A}^{-1} = \begin{bmatrix} A^{-1} - v v^T & \frac{v}{\gamma} \\ \frac{v^T}{\gamma} & -\frac{1}{\gamma} \end{bmatrix}$$

where

$v = A^{-1} w$

$\gamma = w^T v - \alpha$

Update formula #4 (deactivated linear constraint)

Assume now that a primal linear constraint becomes inactive. The corresponding dual variable is set to zero and it is removed from the dual subspace of positive variables. First let us interchange the last positive dual variable with the one set to zero.

Then if the old inverse Hessian was:

$$A^{-1} = \begin{bmatrix} \tilde{A}^{-1} & w \\ w^T & \alpha \end{bmatrix}$$

the new inverse Hessian is updated by the formula:

$$\tilde{A}^{-1} = A^{-1} - ww^T/\alpha$$

Preliminary results indicate that the computational cost for solving a separable quadratic problem is of the same order of magnitude as for inverting a matrix whose dimensionality is equal to the number of active linear constraints. When applied to a non quadratic problem, the method is employed iteratively at the expense of only a few matrix inversions. As an illustration of the power of this dual solver, explicit problems of the CONLIN type, with 300 design variables and 500 constraints, can be solved on a PC within a few hours.

References

[1] Lasdon, L.S. (1970) **Optimization Theory for Large Systems**, Macmillan, New York, 396-459.

[2] Lootsma, F.A. (1989) 'A comparative study of primal and dual approaches for solving separable and partially separable nonlinear optimization problems', **Structural Optimization** 1, 73-79.

[3] Fleury, C. (1979) 'Structural weight optimization by dual methods of convex programming', **International Journal for Numerical Methods in Engineering** 14, 1761-1783.

[4] Fleury, C. (1982) 'Reconciliation of mathematical programming and optimality criteria approaches to structural optimization', Chapter 10 of **Foundations of Structural Optimization: a Unified Approach** (A.J. Morris, ed.), John Wiley and Sons, 363-404.

[5] Venkayya, V.B., Khot, N.S. and Berke, L. (1973) 'Application of optimality criteria approaches to automated design of large practical structures', Proc. **Second Symp. Struct. Opt.**, AGARD-CP-123, Milan, Italy.

[6] Schmit, L.A. and Miura, H. (1976) 'Approximation concepts for efficient structural synthesis', NASA Contractor Report, NASA-CR 2552.

[7] Fleury, C. and Schmit, L.A. (1980) 'Dual methods and approximation concepts in structural synthesis', NASA Contractor Report, NASA-CR 3226.

[8] Braibant, V. and Fleury, C. (1985) 'An approximation concepts approach to shape optimal design', **Computer Methods in Applied Mechanics and Engineering** 53, 119-148.

[9] Fleury, C. (1986) 'Shape optimal design by the convex linearization method', Chapter 12 of **The Optimum Shape: Automated Structural Design** (J. Bennett and M. Botkin, eds.), Plenum Press, 297-326.

[10] Fleury, C. and Braibant, V. (1986) 'Structural optimization - a new dual method using mixed variables', **International Journal for Numerical Methods in Engineering** 23, 409-428.

[11] Fleury, C. (1989) 'CONLIN: an efficient dual optimizer based on convex approximation concepts', **Structural Optimization** 1, 81-89.

[12] Nguyen, V.H., Fleury, C. and Strodiot, J.J. (1987) 'A mathematical convergence analysis of the convex linearization method for engineering design optimization', **Engineering Optimization** 11, 195-216.

[13] Starnes, J.H. and Haftka, R.T. (1979) 'Preliminary design of composite wings for buckling, stress and displacement constraints', **Journal of Aircraft** 16, 564-570.

[14] Svanberg, K. (1987) 'Method of moving asymptotes - a new method for structural optimization', **International Journal for Numerical Methods in Engineering** 24, 359-373.

[15] Smaoui, H., Fleury, C. and Schmit, L.A. (1988) 'Advances in dual algorithms and convex approximation methods', Proc. AIAA/ASME/ASCE 29th Structures, Structural Dynamics, and Materials Conference, 1339-1347.

SEQUENTIAL CONVEX PROGRAMMING FOR STRUCTURAL OPTIMIZATION PROBLEMS

C. FLEURY
Aerospace Laboratory
University of Liege
Rue E. Solvay, 21
B-4000 Liege, Belgium

ABSTRACT. In this Lecture, several recent methods based on convex approximation schemes are discussed, that have demonstrated strong potential for efficient solution of structural optimization problems.

First, the now well established "Approximation Concepts" approach is briefly recalled for sizing as well as shape optimization problems. Next, the "Convex Linearization" method (CONLIN) is described, as well as one of its recent generalization, the "Method of Moving Asymptotes" (MMA). Both CONLIN and MMA can be interpreted as first order convex approximation methods, that attempt to estimate nonlinearity on the basis of semi-empirical rules.

Attention is next directed toward methods that use diagonal second derivatives in order to provide a sound basis for building up high quality explicit approximations of the behaviour constraints. In particular, it is shown how second order information can be effectively used without a prohibitive computational cost.

Various first and second order approaches have been successfully tested on simple problems that can be solved in closed form, on sizing optimization of trusses, and on two-dimensional shape optimal design problems. In most cases convergence is achieved within five to ten structural reanalyses.

1. Introduction

This Lecture describes recent optimization methods particularly well adapted to solve many problems arising in structural design. Mathematically the numerical optimization problem considered herein can be written in the following general form:

$$\text{minimize} \quad c_0(x) \tag{1}$$

$$\text{subject to} \quad c_j(x) \leq 0 \qquad j=1,m \tag{2}$$

Although the objective function (1) is very often a linear function of the design variables x_i (classical weight minimization with sizing variables), herein $c_0(x)$ will be assumed to be possibly a nonlinear function, so that it can represent any structural characteristic to be minimized (e.g. stress concentration), with sizing or shape design variables. The inequalities (2), often called behaviour constraints, impose limitations on structural response quantities such as stresses and displacements under static loading cases. These constraints are usually nonlinear functions, but in some situations, they might also

G. I. N. Rozvany (ed.), Optimization of Large Structural Systems, Vol. I, 531–553.