Completely Automatic Weight-Minimization Method for High-Speed Digital Computers

G. C. Best
General Dynamics, Fort Worth, Tex.

This paper presents a step-by-step procedure, related to the method of steepest descent, for the weight minimization of an arbitrary structure and, as an example, applies it to optimizing the weight of a simple cantilevered box. The procedure presented is capable of handling limiting conditions placed upon the stresses or deflections at selected points and is, in general, applicable to any single-extrema type of continuous function. It is assumed that a procedure or program exists—the stiffness-matrix method being ideal for this purpose—for calculating stresses and deflections, etc., at specific points under various load conditions, once the defining structural parameters are known. Hence, only auxiliary calculations needed to interpret output from and prepare new input for such a program are discussed. Furthermore, the input is assumed to vary continuously, and the subject of discretely varying arguments, as well as multiple extrema, is considered to be outside the scope of the present paper.

Nomenclature

d = vector of gap openings
g = col($\partial w/\partial x_1$, $\partial w/\partial x_2$, . . . $\partial w/\partial x_m$), gradient vector of weight function $w(x)$
p = load vector
u = unit direction vector
$\tilde{u}$ = vector for change in x
v = unnormalized direction vector
w = $w(x, x_2, \ldots x_m)$, weight function
x = col(x_1, x_2, . . . x_m), input vector
δ = deflection vector
σ = col(σ_1, σ_2, . . . σ_i), output vector
$\Delta\sigma$ = col($\Delta\sigma_1$, $\Delta\sigma_2$, . . . $\Delta\sigma_t$), gap vector
σ^b = col($\sigma_1{}^b$, $\sigma_2{}^b$, . . . $\sigma_t{}^b$), output bounds vector
λ = a scalar
$\Delta\lambda$ = step length
A = cap area
C = matrix formed from R columns associated with closed gaps
R = $(r_{ij}) = (\partial\sigma_j/\partial x_i)$, partial derivative matrix
S = stiffness matrix
ΔS = change in S matrix due to small change in an input parameter
Γ = col(γ_1, γ_2, . . . γ_t), γ_i as in Eq. (16)
Λ = col(λ_1, λ_2, . . . λ_r), Lagrange multipliers

Introduction

In the classical type of minimization problem, extrema are determined from the property that all partial derivatives of the function being minimized are zero at such points. However, in typical engineering minimization problems, the minima are usually not determined by this property but rather because limiting values exist which must not be exceeded. Hence, any practical procedure for handling minimization problems in engineering must encompass both types, and this, together with the fact that structures must be satisfactory under many different load conditions, complicates the problem of structural weight minimization considerably. The method presented in this paper, however, excluding multiple minima, provides a feasible solution, at least on high-speed digital or analog-digital computing machines.

A steepest descent step method, described briefly in Ref. 1, is used. Each step is made as large as possible without exceeding some design conditions, linearity being assumed when each step length is calculated. A main (stiffness-matrix) program is assumed to exist with fixed load input, for L conditions, and with variable input x = col(x_1, x_2, . . . x_m), consisting of structural parameters, i.e., length, thicknesses, moduli, etc. (input is subject in most cases to lower bounds, since areas, etc., cannot be negative), and output σ = col(σ_1, σ_2, . . . σ_t) consisting of stresses, deflections, etc., at certain points in the structure (all load conditions are represented simultaneously in this single output vector). In the calculations of the main program, each output variable σ_i is assumed to be a continuous function of the input x_i. To each such output function σ_i, bounds $\sigma_i{}^b$ (upper or lower) will be assigned. If a variable has both, it may appear twice in the output vector. The difference between calculated output and allowable output, as expressed by the assigned bounds, will be described as the gap, defined by the formula

$$\Delta\sigma_i = \epsilon_1{}^i(\sigma_i{}^b - \sigma_i) \tag{1}$$

where $\epsilon_1{}^i = +1$ if $\sigma_i{}^b$ is an upper bound, and $\epsilon_1{}^i = -1$ if $\sigma_i{}^b$ is a lower bound; the signs are arranged so that a negative gap implies an exceeded design condition.

A weight function $w(x) = w(x_1, x_2, \ldots x_m)$ is assumed known and assumed to be a single-valued continuous function of x_i. Also, let u be any vector of unit length with m components. Then for any fixed x and u but for variable λ, the relation

$$\frac{dw}{d\lambda}(x + \lambda u) = \Sigma_i \frac{\partial w}{\partial x_i} u_i = g^\tau u \tag{2}$$

holds; the partials $\partial w/\partial x_i$ are the well-known components of the gradient g of w and the superscript τ denotes transposition. Expression (2) can conveniently be described as the rate of change of w in the u direction; this terminology implies that the components of x are altered in a fixed proportion—the proportion of the components of u, which in three-dimensional space would imply motion along a line through the point x in the direction of u.

After a gap has been closed, subsequent u's are selected, so that further changes in x have no effect upon that gap. This feature of the method is of prime importance, since it permits us to deal with many restrictive conditions at one time.

Partial Derivative Matrix

To begin, some point x for the input must be selected, preferably one which gives some value of w close to the minimum weight so as to reduce subsequent calculations, the location of this point being otherwise immaterial. Then, in the neigh-

Received October 30, 1963; revision received March 23, 1964.
* Senior Structures Engineer.

Table 1 Input vectors for each step

Step	Input at start				Weight	Gaps closed at start
	t_1	t_2	t_3	A		
a1	0.0500	0.0500	0.0500	3.1825	1661.85	0
2	0.045793	0.049579	0.049558	3.1875	1576.01	6
3	0.046287	0.035663	0.040833	3.1859	1551.63	6,7
4	0.046098	0.035760	0.203140	3.1831	1536.06	6,7,8
5	0.049586	0.034029	0.035131	0.00111	1095.75	6,7,8[a]
c5	0.050116	0.033855	0.025025	0	1105.96	...

[a] Gap 4 is critical but cannot be closed, since this would require a negative cap area.

borhood of this point, approximations are made of the partial derivatives $\partial\sigma_j/\partial x_i$, and these are assembled into the $m\chi t$ matrix

$$R = (r_{ij}) \qquad r_{ij} = \partial\sigma_j/\partial x_i \qquad i = 1, 2, \ldots, m \qquad j = 1, 2, \ldots, t \tag{3}$$

These partial derivatives may be conveniently calculated by the "equivalent load" method. Reference 5 gives a closely related procedure. To describe this method briefly, let the stiffness-matrix equation of the structure be

$$S\delta = p \tag{4}$$

where

δ = deflection vector
p = load vector
S = stiffness matrix

Then if a small change Δx_i (equaling, for example, 0.001) is made in an input variables x_i, giving a new input vector x', the change ΔS is the stiffness matrix—usually affecting only a small portion of the over-all S matrix—can be calculated and Eq. (4) rewritten as

$$(S + \Delta S)\bar{\delta} = p \tag{5}$$

where $\bar{\delta}$ is now written for δ, the deflections having necessarily changed and the loads p remaining unaltered. Rearranging Eq. (5) and multiplying through by $F = S^{-1}$ gives

$$\bar{\delta} = F(p - \Delta S\bar{\delta}) \tag{6}$$

which, since $\bar{\delta}$ is not known, can best be solved by the iterative formula

$$\delta^{(\nu+1)} = F(p - \Delta S\delta^{(\nu)} \qquad \nu = 0, 1, 2 \ldots \tag{7}$$

starting with $\delta^{(0)} = 0$. This converges rapidly to $\bar{\delta}$ when ΔS is small, and the main stiffness-matrix program then provides $\sigma(x')$ as soon as $\bar{\delta}$ is known. A satisfactory approximation of the ith column of the partial derivative matrix R is then given by $(\sigma(x') - \sigma(x))/\Delta x_i$.

In the foregoing, a structural change has been effected as though it were a load (hence, the name), thus avoiding inversion of the over-all stiffness matrix each time an R column is calculated. Without this bypass of the S matrix inversion, a time-consuming operation, this minimization procedure would be impractical.

Having obtained the partial derivative matrix R, one can then calculate, as for Eq. (2), the change in any output variable σ_j for a unit change in the direction u, from the relation

$$\frac{d\sigma_j}{d\lambda}(x + \lambda u) = \frac{\partial\sigma_j}{\partial u} = \Sigma_i \frac{\partial\sigma_j}{\partial x_i} u_i = (R_{.j})^T u \tag{8}$$

where $R_{.j}$ is the jth column of the matrix R. The value shown in Eq. (8) constitutes the jth component of the vector $R^T u$, needed in what follows.

Step Direction

Assume that r gaps have been closed and that it is desired to determine a unit vector u as the direction for the next step. If $r = 0$, as it does at the start, the normalized negative of the gradient, i.e., the customary steepest descent, will be used. When $r \neq 0$, it is desirable to choose u so that the gaps already closed will not be altered. Hence, letting C be the $m \times r$ matrix formed from those columns of R associated with the closed gaps, it can be seen from Eq. (8) that the vector u must be chosen so that

$$C^T u = 0 \tag{9}$$

Otherwise one may choose u with the sole restriction that it be a unit vector. Hence it will be chosen to minimize expression (2) and give it a most rapid rate of weight decrease.

Following the Lagrange multiplier method (Ref. 2) one therefore writes

$$\varphi = \sum_i g_i u_i + \sum_i \sum_{k=1}^{r} \lambda_k (C^T)_{ki} u_i + \lambda_0 \left(\sum_i u_i^2 - 1\right) \tag{10}$$

where the λ's are the Lagrange multipliers. Taking partials with respect to u_j, the partials with respect to the λ's merely giving Eq. (9) again, or the requirement that u be a unit vector, one obtains

$$\frac{\partial\varphi}{\partial u_j} = g_j + \sum_{k=1}^{r} \lambda_k (C^T)_{kj} + 2\lambda_0 u_j = 0 \qquad j = 1, 2, \ldots m \tag{11}$$

or in matrix form

$$g + C\Lambda + 2\lambda_0 u = 0 \tag{12}$$

Table 2 Direction vectors and step lengths

Step	Direction vector				Est. step length	Final step length
	u_1	u_2	u_3	u_4		
b1	−0.99451	−0.09945	−0.03165	−0.00781	0.00423	0.00423
2	+0.02964	−0.83505	−0.54146	−0.09287	0.23069	0.01667
3	−0.01054	+0.00539	−0.98754	−0.15693	0.031994	0.01799
4	+0.00110	−0.00054	+0.00063	−0.99999	18.3172	3.1820
5[a]	+0.00053	−0.00017	−0.00008	−0.00111	...	...

[a] $\bar{u}$ is determined from Eq. (17).

Table 3 Partial matrices and gamma vector

		Condition I					Condition II				
Step	Item	τ_1 1	τ_2 2	τ_3 3	σ_4 4	δ_1 5	τ_1 6	τ_2 7	τ_3 8	σ_4 9	δ_1 10
	σ^b	40	40	40	−70	0.030	−40	40	40	−70	0.025
	ϵ_1	1	1	1	−1	1	−1	1	1	−1	1
a1	σ	0	11.783	0	−24.147	0.01083	−37.170	28.548	18.585	−20.415	0.01104
a1	$\Delta\sigma$	40	28.217	40	45.853	0.01917	2.8296	11.452	21.415	49.584	0.01396
a1	R	0	0	0	314.55	−0.12552	669.80	−36.640	35.510	363.72	−0.16896
		0	−235.21	0	10.474	−0.02801	29.421	−555.28	−14.920	−51.501	−0.03561
		0	−0.02	0	0.043	0.00004	−5.422	−2.608	−367.68	11.203	−0.00469
		0	−0.01	0	2.463	−0.00096	0.701	0.331	−0.241	0.310	−0.00011
b1	R^Tu	0	23.402	0	−313.89	0.12762	−668.88	91.742	−22.192	−356.96	0.17152
b1	Γ	∞	1.20575	∞	0.14608	0.15022	0.00423	0.12483	0.96497	0.13891	0.08140
a2	σ	0	11.883	0	−25.540	0.01139	−40.238	28.950	18.480	−22.282	0.01182
a2	$\Delta\sigma$	40	28.117	40	44.460	0.01861	−0.23821	11.050	21.520	47.718	0.01317
a2	R	0	0	0	347.55	−0.13868	772.89	−41.021	40.863	477.43	−0.19464
		0	−234.95	0	11.747	0.02849	30.866	−558.354	−14.158	−51.005	−0.03623
		0	0	0	0	0	−5.45	−2.404	−360.91	10.861	−0.00458
		0	0	0	2.776	−0.00111	0.914	0.557	−0.298	0.685	−0.00021
b2	R^Tu	0	196.20	0	0	19,784.	0	466.32	208.48	50.797	26,984
b2	Γ	∞	14,331.0	∞	∞	0.94043	∞	0.02370	0.10322	0.939	0.48828
a3	σ	0	16.520	0	−25.534	0.01186	−40.370	39.889	22.880	−21.154	0.01248
a3	$\Delta\sigma$	40	23.480	40	44.467	0.01814	−0.3687	0.1165	17.120	48.847	0.01252
a3	R	0.001	0	−0.005	347.40	−0.13860	764.10	−59.400	52.000	469.10	−0.19320
		−0.002	−450.6	−0.744	11.700	−0.05040	54.099	−1053.7	−30.6	−98.600	−0.06740
		−0.002	0	−0.363	0	0	−8.001	−5.200	−542.5	16.3	−0.00700
		0.001	0	−0.489	2.700	−0.00110	0.900	0.500	−0.400	0.700	−0.00020
b3	R^Tu	0.00229	−2.4300	0.43168	−4.0214	0.00136	0	0	535.09	−21.682	0.00862
b3	Γ	16,689.	9.6618	8934.3	11.057	13.324	∞	∞	0.03199	2.2529	1.4528
a4	σ	0	16.475	0	−25.607	0.01189	−40.257	39.954	40.099	−21.770	0.01274
a4	$\Delta\sigma$	40	23.525	40	44.393	0.01811	−0.2567	0.0458	−0.0985	48.230	0.01227
a4	R	0	0	−0.002	349.3	−0.13940	768.60	−56.800	87.50	470.9	−0.19360
		0	−448.2	−0.002	11.8	−0.05010	53.799	−1053.1	−53.6	−97.800	−0.06770
		0	0	−0.002	0	0	−24.401	−15.8	−1637.8	48.900	−0.02110
		0	0	−0.006	2.80	−0.00110	0.799	0.500	−0.900	0.700	−0.00020
b4	R^Tu	0.00166	0.24382	0.00581	−2.4234	0.00084	0	0	0	−0.09967	0.00001127
b4	Γ	22,831	96.434	6764.04	18.317	18.586	∞	∞	∞	483.36	1,082.9
a5	σ	0	17.314	0	−36.493	0.01632	−40.405	39.815	39.862	−21.871	0.01276
a5	$\Delta\sigma$	40	22.686	40	33.507	0.01369	−0.4053	0.1848	0.1381	48.129	0.01224
a5	R	−0.002	0	−0.008	705.5	0.28150	778.5	−15.1	20.4	485.0	−0.1950
		−0.002	−494.3	−0.005	24.0	−0.0596	41.5	−1107.3	−40.9	−112.1	−0.0671
		−0.001	0	−0.009	0	0	−16.10	−11.8	−1510.1	49.4	−0.0209
		−0.002	0	−0.005	5.700	−0.0022	0.900	0.4	−0.9	0.6	−0.0002
5[a]	σ	0	17.4	0	−36.126	0.01618	−39.996	40.006	40.005	−21.596	0.01267

[a] Gaps are closed utilizing Eq. (17), setting $A = 0$.

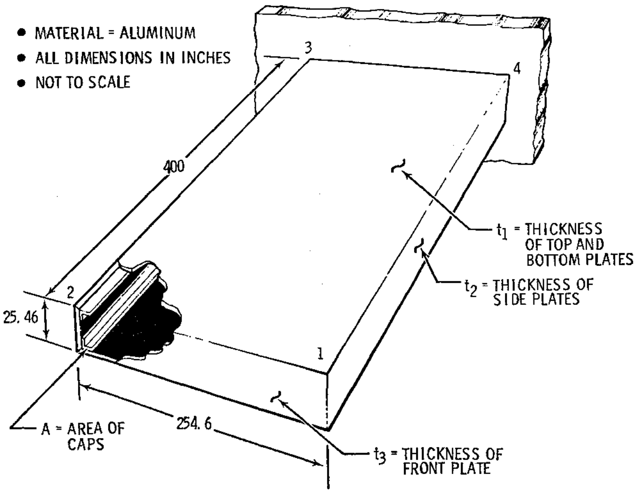

Fig. 1 Cantilevered box.

where $\Lambda = \text{col}(\lambda_1, \lambda_2, \ldots \lambda_r)$. Premultiplying Eq. (12) by C^T and using Eq. (9) gives

$$C^Tg + C^TC\Lambda = 0 \tag{13}$$

The columns of C may be assumed to be linearly independent, since any column which was a linear combination of the others could be dropped without risk of altering the gap involved. Hence C^TC is a nonsingular matrix, and one may premultiply Eq. (13) by $(C^TC)^{-1}$, solve for Λ and insert in Eq. (12), and then solve the resulting equation for u, obtaining

$$u = -\frac{1}{2\lambda_0}[g - C(C^TC)^{-1}C^Tg] = -\frac{v}{2\lambda_0} \tag{14}$$

in which the scalar λ_0 is determined from the fact that the vector u must have unit length. This gives finally

$$u = \pm(v^Tv)^{-1/2}[I - H]g \tag{15}$$

where $H = C(C^TC)^{-1}C^T$. Now, H is an idempotent matrix;[3,4] hence, so is $I - H$. It follows that the latter has only roots 0 or $+1$, and by a well-known theorem[3] the roots of $(v^Tv)^{-1/2}$ $(I \times - H)$ are either 0 or $+(v^Tv)^{-1/2}$, i.e., $\geqslant 0$. Since this matrix is symmetrical, it is positive semidefinite, so when the u with the plus sign in Eq. (15) is used in (2), the expression g^Tu must be either zero or positive. Therefore, if g^Tu is not zero, the extremum with the negative sign in Eq. (15) should be used to give the most rapid decrease in weight per unit step length. If g^Tu is zero, then a classical type of extremum has been reached, indicating, in general, that no further weight reduction is possible.

Table 4 C matrices

Step 2 (6)	Step 3 (6)	Step 3 (7)
772.89	764.10	−59.400
30.866	54.099	−1053.7
−5.45	−8.001	−5.200
0.914	0.900	0.500

Step 4 (6)	(7)	(8)	(R Col.)
768.60	−56.800	87.50	
53.799	−1053.1	−53.6	
−24.401	−15.8	−1637.8	
0.799	0.5	−0.900	

Step 5, Eq. (17)) with $A = 0$ (6)	(7)	(8)
778.5	−15.1	20.4
41.5	−1107.3	−40.9
−16.10	−11.8	−1510.1

Step Length

After some weight-reducing unit direction vector u has been computed, it is necessary to estimate how far to go in that direction before one or more gaps close. For a first approximation, linear relationships are assumed, and the step length is computed from the formula

$$\Delta\lambda = \epsilon_1{}^i\epsilon_2{}^i \min_i \left\{\frac{\Delta\sigma_i}{|\partial\sigma_i/\partial u|}\right\} = \epsilon_1{}^i\epsilon_2{}^i \min_i \gamma_i \tag{16}$$

where $\Delta\sigma_i$ is the gap of Eq. (1), $\partial\sigma_i/\partial u$ is given by Eq. (8), and $\epsilon_2{}^i$ is the sign of $\partial\sigma_i/\partial u$. In scanning for the minimum γ_i the sign of each $\Delta\lambda_i$ should be calculated, and any γ_i for which $\Delta\lambda_i$ is negative while the gap is positive should be ignored. Let $\Gamma = \text{col}(\gamma_1, \gamma_2, \ldots \gamma_l)$ for future reference.

Since, in general, the functions involved will not be linear, the closure of a gap, i.e., its reduction to some sufficiently small value that was previously agreed on, must usually be accomplished by successive approximations—always varying all the input parameters together in the same proportion, i.e., varying the step length in the assigned direction. All calculations are based upon an assumption of linearity, which is, of course, not usually realized, and this may cause some closed gaps to spring open as attempts are made to close new gaps.

Final Gap Closure

The process just described is repeated until as many gaps are closed as possible, i.e., $r = m$, or until a best possible unit direction vector u is found to give $g^Tu = 0$, indicating that a classical type of extremum has been attained. At this point it may appear desirable to close exactly all the approximately closed gaps. This can be done, certainly, by the method already described, but a quicker way may generally be used:

$$C\tilde{u} = -d \tag{17}$$

where d is the vector of gap openings, C is formed, as before, from R columns associated with the gaps involved, and $\tilde{u}$ is the vector, not necessarily unit, of changes to be made in x. When C is square and nonsingular, Eq. (17) is easily solved, although there is no guarantee that weight will be reduced when it is employed; such changes are small, however, when d is small.

This completes the process and gives the minimum weight sought, subject to the limitations imposed herein, i.e., that there is only one minimum and that the input varies continuously.

Example Problem

An example problem, minimizing the weight of a cantilevered box, is given to illustrate the method, since simplicity, rather than optimum weight, has been emphasized. Only four parameters are varied: the skin thickness t_1 of the top and bottom plates, t_2 of the side plates, t_3 of the front plate, and the cap areas A. These vary equally along the box from the cantilevered end, and no attempt has been made to taper them for less weight, as would obviously have been possible under the assumed load conditions.

To label tabulated results, the individual calculations for each step are summarized. Each step starts from some point $x = \text{col}(t_1, t_2, t_3, A)$, with certain gaps $\Delta\sigma_{\nu_1}, \Delta\sigma_{\nu_2}, \ldots \Delta\sigma_{\nu_r}$ already closed.

a) Calculate the weight $w(x)$, the output vector σ, the gap vector $\Delta\sigma$, Eq. (1), and the partial matrix R, Eq. (3).

b) Calculate the vector u, Eqs. (14) and (15), the vectors R^Tu, Γ, and the step length $\Delta\lambda$, Eq. (16), to close gap ν_{r+1} assuming linearity.

c) Calculate the output vector and gap vector at $x + \Delta\lambda u$ to determine if gap ν_{r+1} is closed, and, if not, close by the method of successive approximations, varying $\Delta\lambda$.

The weight function w for this problem is

$$w = 20368t_1 + 2036.8t_2 + 648.2116t_3 + 160.0A \quad (18)$$

giving a normalized gradient of

$$g^* = \text{col}\ (0.99451, 0.09945, 0.03165, 0.00781) \quad (19)$$

the negative of which will be used as the first unit direction vector u, the starting x being taken as (0.05, 0.05, 0.05, 3.1825). Two load conditions are applied. In condition I a 15,000-lb up-load is applied at points 1 and 2 (see Fig. 1), whereas in condition II a 60,000-lb up-load is applied at point 1 and a similar down-load at point 2. The output consists of shear stresses τ_1, τ_2, τ_3 in the top, side, and front plates, respectively, compressive stress σ_4 in the top cap at point 4, and the deflection δ_1 at point 1. Allowable stresses of 40,000 psi shear and $-70,000$ psi compression were assumed. The deflection at point 1 was limited in condition I to 30 in. and in condition II to 25 in. Hence, the output vector has 10 components, condition I being shown first. Output components are shown $\times 10^{-3}$, and bound and gap values are shown similarly. The problem requires four regular steps, a fifth step, using Eq. (17), being employed to close exactly all the critical gaps. The input and direction vectors for each step are shown in Tables 1 and 2, respectively.

The restriction that the input parameters cannot become negative is encountered in this problem, and, as a consequence, the cap area is set to zero. The output vector for each step, together with the partial matrices R and the vectors $R^T u$ and Γ are shown in Table 3. Table 4 shows the C matrices of Eq. (9).

References

[1] Best, G. C., "A method of structural weight minimization suitable for high-speed digital computers," AIAA J. **1**, 478–479 (1963).

[2] Lanczos, C., *The Variational Principles of Mechanics* (Univ. of Toronto Press, Toronto, Canada, 1949), p. 43.

[3] Perlis, S., *Theory of Matrices* (Addison-Wesley Press, Inc., Reading, Mass., 1952), p. 160, 168.

[4] Householder, A. S., *Principles of Numerical Analysis* (McGraw-Hill Book Co., Inc., New York, 1953), p. 30.

[5] MacNeal, R. H., "Application of the compensation theorem to the modification of redundant structures," J. Aerospace Sci. **20**, 726–727 (1953).