NASA Contractor Report 2991

Efficiency of Unconstrained Minimization Techniques in Nonlinear Analysis

Manohar P. Kamat and Norman F. Knight, Jr.

GRANT NGR-47-004-114
MAY 1978

NASA

NASA Contractor Report 2991

Efficiency of Unconstrained Minimization Techniques in Nonlinear Analysis

Manohar P. Kamat and Norman F. Knight, Jr.
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Prepared for
Langley Research Center
under Grant NGR 47-004-114

National Aeronautics
and Space Administration

Scientific and Technical
Information Office

1978

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

The prediction of transient linear or nonlinear response of structures is almost invariably accomplished by using a temporal finite-difference scheme to effectively eliminate time as a variable and reduce the system to a set of algebraic equations in the unknown nodal variables of the finite element discretization. Finite differencing in time may be either of the explicit or implicit type. Furthermore, the resulting algebraic equations may be at times linearized within a time step. The type of temporal scheme chosen together with response approximations determines the computational effort required to solve for the unknowns at time $(t+\Delta t)$ knowing the same at time t. With an explicit scheme such effort is small although stability considerations limit the size of the time step. On the other hand, with an implicit scheme the computational effort within a time step may be significant but inherent stability of the scheme permits larger time steps compared to those allowed by explicit schemes. Of interest in this paper are those schemes which essentially linearize response within a time step and use an explicit scheme as in DYCAST [1] and those which do not linearize response within a time step and use an implicit scheme as in ACTION [2] .

For schemes which do not linearize the response within a time step, several different techniques for the solution of the nonlinear equations may be used. Such techniques have been discussed at great lengths by Bergan [3] and Stricklin et al [4] . Of particular interest is the technique utilizing the minimization algorithms of mathematical programming. This approach has been used successfully for nonlinear structural analyses

[5]-[7]. In this case, the problem of finding the solution of the equilibrium equations can be equivalently posed as the one corresponding to the minimum value of a potential function. For all structural problems with geometric and material nonlinearities of the type considered herein such a potential function always exists. However, only positive or negative definite systems can be handled by this technique with the result that some unstable equilibrium configurations for which the potential function has a non-definite form remain intractable. This is not a very serious limitation however, since the determination of such equilibrium configurations is usually of academic interest.

The two simultators DYCAST and ACTION are intended for solving inertial problems such as those involving the elastic-plastic large deformation response of structures. It is well-known that, more often than not, such problems are best solved by implicit techniques [8]. Since, the solution effort within a time step with an implicit scheme, with or without linearization, is by no means minimal but presumably comparable to that using minimization algorithms, the latter are appealing in that they solve the actual nonlinear problem. Hence, a rather rigorous evaluation of their effectiveness in predicting nonlinear structural response is deemed necessary. Of course, similar investigations have been carried out by previous investigators [9],[10],[11], but only as regards the minimization of nonlinear mathematical functions. Geometric and material nonlinearities cannot in general be described by smooth nonlinear functions of the type considered by previous investigators. An energy evaluation for an inelastic member is several times more expensive than that for a purely elastic member. Furthermore, the computational effort involved in the function and gradient evaluations in the case of finite element models, with rather specialized connectivity properties leading to banded stiffness matrices, may bear no direct and simple relationship with the computational effort for simple mathematical functions. The conclusions regarding the effectiveness of these algorithms for solving nonlinear problems of structural analyses are likely to be different. The effectiveness of these algorithms may be very much a function of the type of nonlinearity, geometric or material, and also whether a static or a dynamic problem is being considered.

We have, in this study, examined a host of unconstrained minimization algorithms with a view of determining their effectiveness for solving static or dynamic structural problems involving either (i) geometric, (ii) material, or (iii) a combination of geometric and material nonlinearities.

2. Classification of Algorithms

In the realm of mathematical programming, the algorithms used for unconstrained minimization can be broadly classified into three distinct classes stemming from the level of computational sophistication: (i) the zeroth order requiring only function evaluations: (ii) the first order requiring evaluation of the gradient as well as the function and (iii) the second order requiring, in addition, a variable metric related to the curvatures of the function. Only the techniques belonging to the latter two categories have been more frequently used for structural analysis, apparently because of their higher effectiveness in comparison with zeroth order techniques. The minimization algorithms of the zeroth and first order have very modest storage requirements since at no time is the assembled stiffness matrix of the structure generated. For linear problems with analytic derivatives (derivatives not evaluated through finite differencing), some of these techniques can be nearly as efficient as the conventional stiffness matrix formulation coupled with a Choleski decomposition technique.

The following algorithms, details of which can be found in Appendix A, were examined for their effectiveness in solving static and dynamic structural problems involving geometric and/or material nonlinearities:

1. Zeroth Order Algorithms
   a. Nelder and Mead's Simplex Method [12]
   b. Powell's Conjugate Direction Method [13]
2. First Order Algorithms
   a. Method of Steepest Descent
   b. Fletcher-Reeves' Conjugate Gradient Algorithm [14]
   c. Jacobson-Oksman's Algorithm [15]
3. Second Order Algorithms
   a. Davidon's Variance Algorithm [16]
   b. Davidon-Fletcher-Powell's Variable Metric Algorithm [17,18]
   c. Fletcher's New Variable Metric Algorithm [19]

The efficiency of any given algorithm is quantified by the number of function and gradient evaluations (or equivalent function evaluations) and by the amount of total CPU time required to solve a given problem to within a prescribed degree of accuracy. Previous investigators have also utilized the number of minimizations as a criterion for efficiency. However, it is felt that this alone is not a meaningful criterion because the computational effort involved in a single minimization of a given technique may be significantly different from that of another. Thus, the total number of minimizations is not a meaningful measure of total computational effort. Both analytic and finite difference derivatives are used with a view to determining the sensitivity of these algorithms to inaccuracies in gradient evaluations and to determine the computational efficiency ensuing from the use of an analytic gradient. Development of an analytic gradient presents no insurmountable problems when only geometric nonlinearities exist. With material nonlinearities however, the picture is fundamentally different because the bookkeeping effort increases significantly. The details of the calculation of analytic derivatives for nonlinear structural problems of the type considered herein are presented in reference [20].

3. Classification of Problems and Their Formulation

The problems to be analyzed are arranged into five different categories stemming from the type or degree of nonlinearity and from whether the problem is static or dynamic. Four categories of the static variety were chosen and include the following: (i) the elastica problem, Figure 1, restricted to mild geometric nonlinearities, $P/P_{cr}<0.4$; (ii) the rod-spring problem, Figure 2, which is geometrically highly nonlinear; (iii) an elastic-plastic cantilever beam subjected to a tip bending moment, Figure 3, with only material nonlinearties (deformations and rotations constrained to be small and geometrically linear); and

(iv) an elastic-plastic cantilever beam subjected to a tip bending moment with rotations allowed to be moderately large. The fifth (v) category is a transient problem: an impulsively loaded elastic-plastic clamped beam, Figure 4, wherein the rotations are allowed to be moderately large. The structural response of each system as predicted in this study is in excellent agreement with that presented by Huddleston [21] for the elastica problem; by Haisler, Stricklin and Stebbins [22] for the rod-spring problem; and by Belytschko and Schoeberle [23] for the impulsively loaded elastic-plastic clamped beam problem.

The minimization approach as applied to the solution of transient nonlinear structural problems consists of minimizing a potential function associated with the system for an assumed relationship between displacements and time. The solution process accommodates static response of massless systems and transient response of systems wherein some of the components may have zero masses as special cases. For the purposes of this study, the displacement-time relation for each generalized nodal variable of our finite element model is assumed to be of the form [24]

$$X_{ei} = a_i + b_i\,(\Delta t) + \frac{c_i}{2}\,(\Delta t)^2 + \frac{\beta_i}{6}\,(\Delta t)^3 \qquad (1)$$

where X_{ei} is the i-th generalized nodal displacement at the end of the time step and a_i, b_i, c_i and β_i are constants. These constants are determined in terms of generalized displacement, X_{oi}, velocity, $\dot{X}_{oi}$, and accelerations, $\ddot{X}_{oi}$, at the beginning of the time step and the generalized displacement, X_{ei}, at the end of the time step. The equation of motion corresponding to the i-th degree of freedom takes the form

$$M_i\left[\frac{6}{(\Delta t)^2}(X_{ei}-X_{oi})-\frac{6}{(\Delta t)}\dot{X}_{oi}-2\ddot{X}_{oi}\right]-F_i\Big|_{(t_o+\Delta t)}+\frac{\partial U}{\partial X_{ei}}=0 \qquad (2)$$

where M_i, F_i and U are respectively, the mass, exciting force and strain energy corresponding to the i-th degree of freedom. The equilibrium equations, Eq. (2) for $i=1,\ldots,N$ for a system with N degrees of freedom, can be regarded as the stationary conditions for the functional S

$$S=\sum_{i=1}^{N}\left\{\left[\frac{3}{(\Delta t)^2}X_{ei}^2-\left(\frac{6}{(\Delta t)^2}X_{oi}+\frac{6}{(\Delta t)}\dot{X}_{oi}+2\ddot{X}_{oi}\right)X_{ei}\right]M_i - F_i\Big|_{(t_o+\Delta t)}X_{ei}))\right\}+U+C \qquad (3)$$

where C is an arbitrary constant. Because the masses, M_i, $i=1,\ldots,N$ are positive quantities and U is a positive semi-definite function for most structural materials, the space S can be shown to be convex. Equations (2), for $i=1,\ldots,N$ are thus necessary conditions for S to be a minimum.

Once the assumption of the displacement-time relation is made, the minimization approach, unlike the incremental stiffness approach, solves the actual nonlinear problem within a given load or time step without linearization. Consequently, iteration at constant load to improve the equilibrium or force imbalance at the end of a load or time step is not required.

4. Discussion of Results

The performance of the selected unconstrained minimization algorithms will now be summarized with regard to their effectiveness in predicting nonlinear structural response. However, in order to sanction an inter-class as well as an intra-class assessment of the zeroth, first and second order algorithms, a factor which is a composite of the number of function and gradient evaluations was defined and termed the number of equivalent member energy evaluations. The determination of this factor depends on how the gradient is calculated. The i-th component of the gradient of S using central differences is given by

$$\frac{\partial S}{\partial X_{ei}} = \frac{S(t, X_{e1}, X_{e2} \ldots X_{ei} + \Delta X_{ei}, X_{ei+1} \ldots X_{eN}) - S(t, X_{e1}, X_{e2} \ldots X_{ei} - \Delta X_{ei}, X_{ei+1} \ldots X_{eN})}{2\,\Delta X_{ei}} \quad (4)$$

where ΔX_{ei} is a small change in the i-th component. In computing S in equation (4), only the energies of the affected members are recomputed. Thus, in using central difference operation, each component of the gradient vector corresponding to any one nodal degree of freedom involves at least two member energy evaluations, four if the node is common to two elements, six if the node is common to three elements, and so on. For an analytic gradient, however, the gradient calculation involves only a single function which is similar to the function for the member energy evaluation (see reference 20 for details). The i-th component of the gradient vector may be written as

$$\frac{\partial S}{\partial X_{ei}} = M_i \ddot{X}_{ei} - F_i + \frac{\partial U}{\partial X_{ei}} \quad (5)$$

The term requiring an analytic expression is $(\partial U/\partial X_{ei})$ which can be evaluated as

$$\frac{\partial U}{\partial X_{ei}} = \sum_{k=1}^{m} \int_{v_k} \frac{\partial W}{\partial X_{ei}} dv_k = \sum_{k=1}^{m} \int_{v_k} \left(\frac{dW}{d\varepsilon}\right)_k \left(\frac{\partial \varepsilon}{\partial X_{ei}}\right)_k dv_k$$

or equivalently for one step incremental loading or unloading

$$\frac{\partial U}{\partial X_{ei}} = \sum_{k=1}^{m} \int_{v_k} \sigma_k \left(\frac{\partial \varepsilon}{\partial X_{ei}}\right)_k dv_k \tag{6}$$

where W represents the strain energy density and m is the number of members or elements which has the i-th degree of freedom in common. Thus, the term $\frac{\partial U}{\partial X_{ei}}$ involves a volume integral which is very similar to that required for a member energy evaluation. Hence, it may be assumed that each component of the analytic gradient vector involves approximately at least one member energy evaluation, two if the node is common to two elements, three if the node is common to three elements and so on. Consequently, a significant reduction in the number of equivalent member energy evaluations and in CPU time should be realized if analytic gradients are used instead of finite difference gradients.

4.1 Results Using Finite Difference Gradients

For most minimization algorithms using finite difference approximations for the gradient, the computational effort required to evaluate the gradient using a central difference operator may be nearly twice that required for either a forward or backward difference operator. This computational advantage is partially offset by the fact that the accuracy of the derivatives obtained by utilizing the forward or backward difference operator is likely to be poor. Our limited studies indicate that for first order methods considerable savings in computational time may be realized for some problems by using forward or backward difference derivatives. It is tacitly assumed however that for the calculation of such derivatives, stepsizes are much smaller (approximately 1/100th to 1/1000th; with the minimum being dictated by the precision of the computer) than the stepsizes used for central difference derivatives. The same computational advantage may not be realized with second order methods utilizing a variable metric which is updated recursively using current function and gradient information and converges to the inverse of the Hessian matrix of the functional S. In this case, the inaccuracies in the variable metric resulting from forward or backward difference derivatives can retard the convergence rate substantially, thus offsetting the savings accrued from the relatively cheaper gradient evaluations. For second order methods, the only time forward or backward difference derivatives appear to be competitive with central difference derivatives is when the initial variable metric, $\underline{H}_0$, is set equal to the identity matric, $\underline{I}$, at the beginning of each load or time step of a nonlinear analysis and is not carried forward as the initial variable metric of the next step. Our limited experiments using forward and backward difference derivatives have revealed that the performance of most minimization algorithms employing such derivatives is unpredictable and quite sensitive to the type and degree of nonlinearity in question, even more so than with central difference derivatives. Hence, based on this study, it is recommended that their use should be avoided whenever possible.

The computational effort expended by the various minimization algorithms using either analytic or central difference gradients is documented in the following tables by three weighting factors normalized to their respective minimum value; such that, the most efficient algorithm will have weighting factors of 1.0. The first of these corresponds to the number of minimizations, the second to the number of equivalent member energy evaluations and the third to the elapsed CPU time. For the results reported in Tables 1 through 5, all gradients were evaluated using a central difference operator.

The effects of geometric nonlinearities on the computational effort required by the various algorithms can be observed in Table 1 for a mildly nonlinear example and in Table 2 for a highly nonlinear example While none of the first order methods considered were competitive, Powell's conjugate direction method presents stiff competition to second order methods which set the initial variable metric to the identity matrix $\underline{H}_0 = \underline{I}$ at the beginning of each load step. for the degree of geometric nonlinearity considered, Davidon-Fletcher-Powell's method using an updated variable metric ($\underline{H}_0 = \underline{I}$ only initially) is more efficient. Also, the computational advantage of second order methods which update the variable metric recursively over those which do not, can be readily seen in Table 1. The extremely poor performance of some of the methods in Table 1 suggest that they hold little promise for other cases of nonlinearities.

The effects of material nonlinearities arising from loading and unloading while in the inelastic range can be seen in Table 3. For this study, a linear elastic response is considered for the first two loading steps followed by two loading steps in the inelastic range and a final unloading step in the elastic range. For the linear elastic portion where equation (3) has a quadratic form, all methods performed at least marginally well with the exception of Nelder-Mead's method and the method of steepest descent which performed poorly. An unexpected result came from the fact that the Jacobson-Oksman algorithm had to be restarted repeatedly to ensure linear independence of the search directions. This is reported to have never been necessary by Jacobson and Oksman [15] for their test functions. Although initially superior for the linear elastic steps, Davidon-Fletcher-Powell's method is less competitive than Powell's conjugate direction method, (a zeroth order method) for inelastic loading. Fletcher's method is initially very sluggish in finding a solution but is found to be competive after the first load step and superior during inelastic loading. However the superiority of Davidon-Fletcher-Powell's method is regained when unloading occurs and overall it is again the more efficient method. This is as expected in that methods, like Davidon-Fletcher-Powell's method, which utilize the property of quadratic termination will be more efficient for problems involving quadratic functionals than methods like Fletcher's method which discard this property.

Table 4 summarizes the computational effort expended in solving a problem with a high degree of geometric and material nonlinearities. Only Fletcher's method was successful in solving this problem within a reasonable amount of computer time for the specified load steps as

Davidon-Fletcher-Powell's method failed to converge during the first step in the inelastic range. The reason for this failure to converge could be attributed to either taking too large of a load step or to inaccuracies in the variable metric resulting from the use of central difference gradients.

Table 5 represents the results from a transient problem wherein, to a moderate degree, both geometric and material nonlinearities were included. Again overall superiority was demonstrated by Fletcher's method using an updated variable metric.

4.2 Results Using Analytic Gradients

The above results, together with our experiments with other finite difference operators, clearly indicate that all first and second order algorithms are critically influenced by the accuracy of the gradient vector. This leads very naturally to the consideration of an analytically derived gradient vector and of its effectiveness. The performance of the first and second order methods using analytic gradients is reported in Tables 6 through 8.

The effects of geometric nonlinearities on the computational effort required can be observed in Table 6 for a mildly nonlinear problem and in Table 7 for a highly nonlinear problem. These results again indicate the superior performance of the second order methods. Also, as the degree of geometric nonlinearity increases or in other words as Eq. (3) becomes more nonquadratic, the superiority of Fletcher's method over Davidon-Fletcher-Powell's method becomes more and more evident. It should also be noted that Fletcher-Reeves' method using analytic gradients was able to converge for all load steps considered whereas using central difference gradients, it was not.

The effects of material nonlinearities are reported in Table 8. For this study, two elastic and two inelastic loading steps are made. For the linear elastic portion, the results correlate well with those reported in Table 3. However, for inelastic loading, Davidon-Fletcher-Powell's method maintained its superiority over Fletcher's method. This is quite interesting in that a definite change in performance is observed from that reported in table 3.

The relative performance of the methods using central difference and analytic gradients is documented in Tables 9-11. Analytically derived gradients markedly improve the performance of the first and second order methods with the most impressive improvement occurring when only material nonlinearities are considered, Table 11. It is to be expected that as the number of inelastic members increases, the relative performance of any first or second order algorithm utilizing analytic gradients will continue to improve.

5. Conclusions

The performance of the selected unconstrained minimization algorithms has been summarized with regards to their effectiveness in predicting nonlinear structural response. Based on the data reported, Powell's conjugate direction method is the only zeroth order method which demonstrated good performance, particularly when only material nonlinearities were considered. None of the first order methods considered were competitive and in general, their use is not recommended for nonlinear structural analysis. Despite inaccuracies in the variable metric due to central difference derivatives, the second order method are the more efficient algorithms, primarily because they exploit the curvature related information of the function. Davidon's variance algorithm is not recommended for general structural analysis because of its critical dependence on three user supplied parameters. When mild geometric or material nonlinearities are considered separately, Davidon-Fletcher-Powell's method is found to be more cost-effective on an overall basis over the range of load or time steps considered herein. However, it is found that overall Fletcher's method surpasses Davidon-Fletcher-Powell's method when geometric and material nonlinearities are combined for both static and dynamic problems. Furthermore, the results indicate that with time and an increasing degree of geometric nonlinearity, Fletcher's method becomes increasingly cost-effective.

The sensitivity of the first and second order algorithms to numerical differentiation and the computational efficacy ensuing from the use of analytic gradients was also determined. The use of analytic gradients results in a substantial savings in computational effort. This saving is the result of not only a cheaper gradient evaluation but in most cases, a faster convergence to the solution because of the higher accuracy of all computed quantities.

Therefore, the following recommendations can be made. For general nonlinear structural analysis, Fletcher's new variable metric method using an updated variable metric and, if possible, analytic gradients is the best minimization algorithm of those considered herein. Since Fletcher's method is initially slow in converging to a solution, it is recommended that further studies centered around a more appropriate choice for the initial variable metric are needed. If the degree of nonlinearity is mild and only one type of nonlinearity is considered, Davidon-Fletcher-Powell's method using an updated variable metric and, if possible, analytic gradients is preferred. If analytic gradients are unavailable, then central difference gradients should be used. The use of either forward or backward difference should be avoided.

6. References

1. Armen, H., Levine, H., Pifko, A. and Levy, A., "Nonlinear Analysis of Structures, NASA CR-2351, 1974.

2. Kamat, M. P., Melosh, R. J., Swift, G. W., and Killian, D. E., "Crash Safety of General Aviation Aircraft," VPI-E-78 8, Feb. 1978.

3. Bergen, P. G. and Soreide, T., "A Comparative Study of Different Numerical Solution Techniques as Applied to a Nonlinear Structural Problem," Computer Methods in Applied Mechanics and Engineering, 2, 185 (1973).

4. Stricklin, J. A. and Haisler, W. E., "Formulations and Solution Procedures for Nonlinear Structural Analysis," Computers and Structures, 7, 125 (1977).

5. Bogner, F. K., Mallett, R. H., Minich, M. D. and Schmit, L. A., "Development and Evaluation of Energy Search Methods of Nonlinear Structural Analysis," AFFDL-TR-65-113, WPAFB, Dayton, Ohio, 1966.

6. Young, J. W., "CRASH: A Computer Simulator of Nonlinear Transient Response of Structures," DOT-HS-091-1-125-13, March 1972.

7. Mallett, R. H. and Berke, L., "Automated Method for Large Deflection and Instability Analysis of Three Dimensional Truss and Frame Assemblies," AFFDL-TR-66-102, WPAFB, Dayton, Ohio, 1966.

8. Belytschko, T., Osias, J. R. and Marcal, P. V., "Explicit Integration - Stability, Solution Properties, Cost," Winter Annual Meeting of the ASME, AMD-14, 1 (December, 1975).

9. Fletcher, R., "Function Minimization Without Evaluating Derivatives - A Review," The Computer Journal, 8, 33 (1965).

10. Powell, M. J. D., "A Survey of Numerical Methods for Unconstrained Optimization," SIAM Review, 12 (1), 79 (1970).

11. Straeter, T. A. and Hogge, J. E., "A Comparison of Gradient Dependent Techniques for the Minimization of an Unconstrained Function of Several Variables," AIAA Journal, 8 (12), 2226 (1970).

12. Nelder, J. A. and Mead, R., "A Simplex Method for Function Minimization," The Computer Journal, 7, 308 (1965).

13. Powell, M. J. D., "An Efficient Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivatives," The Computer Journal, 7, 155 (1964).

14. Fletcher, R. and Reeves, C. M., "Function Minimization by Conjugate Gradients," The Computer Journal, 7, 149 (1964).

15. Jacobson, D. H. and Oksman, W., "An Algorithm that Minimizes Homogeneous Functions of N Variables in N+2 Iterations and Rapidly Minimizes General Functions," Journal of Mathematical Analysis and Applications, 38, 535 (1972).

16. Davidon, W. C., "Variance Algorithm for Minimization," The Computer Journal, 10, 406 (1968)

17. Fletcher, R. and Powell, M. J. D., "A Rapidly Convergent Descent Method for Minimization," The Computer Journal, 6, 163 (1963).

18. Stewart, G. W., "A Modification of Davidon's Minimization Method to Accept Difference Approximations of Derivatives," Journal of the Association for Computing Machinery, 14 (1), 72 (1967).

19. Fletcher, R., "A New Approach to Variable Metric Algorithms," The Computer Journal, 13, 317 (1970).

20. Knight, Jr., N. F., "An Efficiency Assessment of Selected Unconstrained Minimization Techniques as Applied to Nonlinear Structural Analyses," M.S. Thesis, Virginia Polytechnic Institute and State University, August 1977.

21. Huddleston, J. V., "A Numerical Technique for Elastica Problems," Journal of the Engineering Mechanics Division, ASCE, 94 (EM5), 1159 (1968).

22. Haisler, W. E., Stricklin, J. A. and Stebbins, F. J., "Development and Evaluation of Solution Procedures for Geometrically Nonlinear Structural Analysis," AIAA Journal, 10 (3), 264 (1972).

23. Belytschko, T. and Schoeberle, D. F., "On the Unconditional Stability of an Implicit Algorithm for Nonlinear Structural Dynamics," Journal of Applied Mechanics, 42, 865 (1975).

24. Newmark, N. M., "A Method of Computation for Structural Dynamics," Journal of the Engineering Mechanics Division, ASCE, 85 (EMS), 67 (1959).

25. Himsworth, F. R., Spendley, W. and Hext, G. R., "The Sequential Application of Simplex Designs in Optimization and Evolutionary Operations," Technometrics, 4, 441 (1962).

26. Davidon, W. C., "Variable Metric Method for Minimization," Argonne National Laboratory, ANL-5990 (Rev.), University of Chicago, 1959.

27. Walsh, G. R., Methods of Optimization, John Wiley and Sons, London, England, 1975.

28. Fox, R. L., Optimization Methods for Engineering Design, Addison-Wesley, Reading, Massachussetts, 1971.

29. Myers, G. E., "Properties of the Conjugate Gradient and Davidon Methods," Journal of Optimization Theory and Applications, 2 (4), 209 (1968).

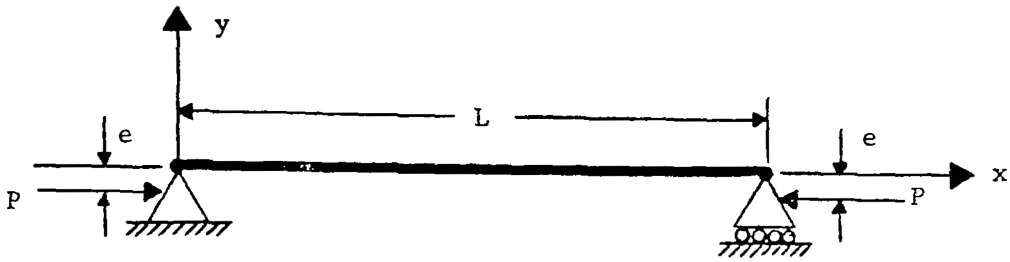
30. Cauchy, A., "Méthode Générale Pour La Résolution Des Systemes D'Equations Simulatanées," Compt. Rend. l'Académie des Sciences, 25, 536 (1847).

31. Davidon, W. C., "Variance Algorithms for Minimization," in Optimization, (R. Fletcher, editor), Chapter 1, Academic Press, London, 1969.

32. Fletcher, R., "A Review of Methods for Unconstrained Optimization," in Optimization, (R. Fletcher, editor), Chapter 1, Academic Press, London, 1969.

33. Crocket, J. B. and Chernoff, R., "Gradient Methods of Maximization," Pacific Journal of Mathematics, 5, 33 (1955).

34. Hestenes, M. R. and Stiefel, E., "Method of Conjugate Gradients for Solving Linear Systems," Journal of Research, National Bureau of Standards, Section B, (49), 409, (1952).



$E_{elas} = 1.958 \times 10^{11}$ Pa

$A = 5.0 \times 10^{-5} m^2$

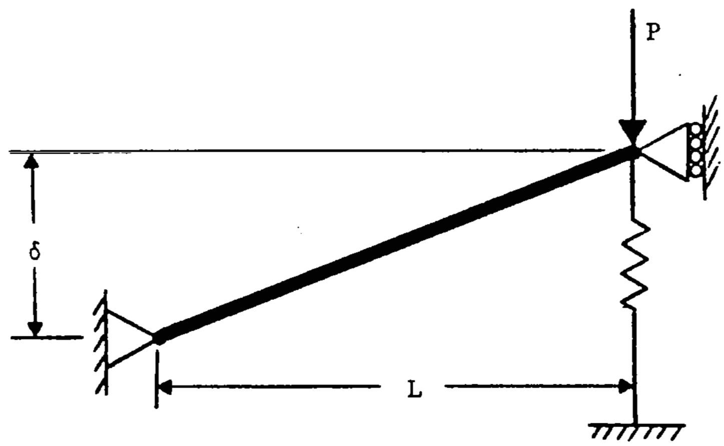
$I_{yy} = 1.042 \times 10^{-10} m^4$

$e/L = 0.01$

$L = 1.0m$

$P_{cr} = 2.013 \times 10^2 N$

Figure 1. Elastica Problem (10 elements, 30 degrees of freedom).



$AE = 4.45 \times 10^7 N$

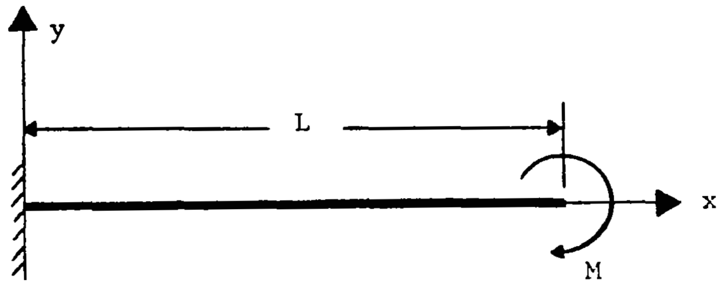
$L = 2.54m$

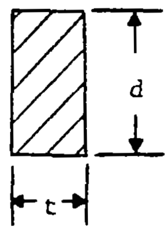
$\delta = 2.54cm$

$\Delta P = 4.45N$

$k_s = 1050N/m$

Figure 2. Rod-Spring Problem (5 elements, 12 degrees of freedom.)





$E_{elas} = 1.958 \times 10^{11}$ Pa

$E_{plas} = 3.751 \times 10^{8}$ Pa

$\sigma_{yp} = 5.585 \times 10^{8}$ Pa

$d = 1.0$cm

$t = 0.5$cm

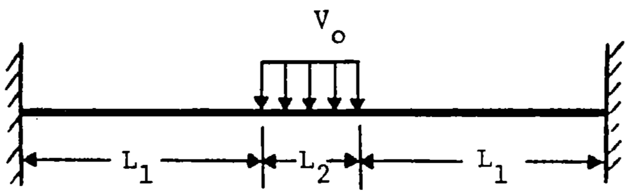
$A = 5.0 \times 10^{-5} m^2$

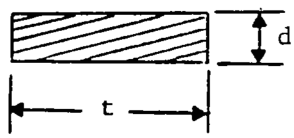
$L = 1.0$m

$I_{yy} = 4.167 \times 10^{-10} m^4$

$M_{yp} = 46.542$ J

Figure 3. Elastic-Plastic Cantilever Beam (4 elements, 12 degrees of freedom, solid rectangular cross section).





L_1 = 127mm

L_2 = 51mm

V_o = 51.18m/sec.

ρ = 7870kg/m^3

E_{elas} = 1.958x10^{11}Pa

E_{plas} = 3.751x10^{8}Pa

σ_{yp} = 5.585x10^{8}Pa

d = 3.18mm

t = 25.40mm

Figure 4.. Impulsively Loaded Elastic-Plastic Clamped Beam (10 elements over half the span, 28 degrees of freedom, solid rectangular cross section).

Table 1.- Performance Data for the Elastica Problem Using Central Difference Gradients.

MINIMIZATION ALGORITHM		APPLIED AXIAL LOAD, P/P_{cr} 0.1	0.2	0.3	0.4
Nelder-Mead's Simplex Method	(m)	91.53*	-	-	-
	(e)	4.42	-	-	-
	(t)	3.99	-	-	-
Powell's Conjugate Direction Method	(m)	1.00	7.91	7.00	6.14
	(e)	2.38	36.78	32.12	28.13
	(t)	2.05	27.80	25.61	21.82
Fletcher-Reeves' Method with restarts after the first N+2 iterations only.	(m)	3.14	24.45	45.83*	-
	(e)	2.30	31.13	55.51	-
	(t)	1.96	23.40	42.06	-
Method of Steepest Descent	(m)	2.54	17.91	17.58**	-
	(e)	1.82	22.78	20.90	-
	(t)	1.47	17.15	16.72	-
Jacobson-Oksman's Method	(m)	1.76	33.09	28.58*	-
	(e)	1.71	69.17	60.21	-
	(t)	1.45	55.03	50.39	-
Davidon's Variance Method with $\underline{H}_0=\underline{I}$	(m)	3.39*	-	-	-
	(e)	2.74	-	-	-
	(t)	1.51	-	-	-
Davidon's Variance Method with $\underline{H}_0=\kappa\underline{I}$ where $\kappa=0.1$	(m)	3.22	17.27	26.00*	-
	(e)	1.51	16.00	23.41	-
	(t)	1.46	13.42	20.61	-
Fletcher's Method with $\underline{H}_0=\underline{I}$ always	(m)	3.78	40.73	31.08	28.14
	(e)	1.77	37.50	28.12	25.56
	(t)	1.59	30.64	23.45	20.62
Fletcher's Method with $\underline{H}_0=\underline{I}$ only initially.	(m)	3.78	1.00	2.00	2.00
	(e)	1.79	1.00	1.87	1.85
	(t)	1.79	1.00	1.72	1.72
Davidon-Fletcher-Powell's Method with $\underline{H}_0=\underline{I}$ always	(m)	1.53	9.18	13.58	11.21
	(e)	1.00	11.82	17.04	14.21
	(t)	1.04	11.44	15.49	12.84
Davidon-Fletcher-Powell's Method with $\underline{H}_0=\underline{I}$ only initially.	(m)	1.53	1.64	1.00	1.00
	(e)	1.00	1.91	1.00	1.00
	(t)	1.00	1.79	1.00	1.00

*Failed to converge for given load increment

**Converged but analysis terminated

(m)...Weighting factor for the number of minimizations.
(e)...Weighting factor for the number of equivalent member energy evaluations.
(t)...Weighting factor for the elapsed CPU time..

Table 2.- Performance Data for the Rod-Spring Problem Using Central Difference Gradients.

MINIMIZATION ALGORITHM		RESULTS FROM THIRTY (30) LOAD STEPS
Powell's Conjugate Method	(m)	3.08
	(e)	7.02
	(t)	5.40
Fletcher's Method with $\underline{H}_0=\underline{I}$ only initially	(m)	1.57
	(e)	1.47
	(t)	1.34
Davidon-Fletcher-Powell's Method with $\underline{H}_0=\underline{I}$ only initially	(m)	1.00
	(e)	1.00
	(t)	1.00

(m)...Weighting factor for the number of minimizations.
(e)...Weighting factor for the number of equivalent member energy evaluations.
(t)...Weighting factor for the elapsed CPU time.

Table 3.- Performance Data for an Elastic-Plastic Cantilever Beam Subjected to a Tip Bending Moment with Small and Geometrically Linear Deformations and Rotations Using Central Difference Gradients.

MINIMIZATION ALGORITHM		APPLIED TIP BENDING MOMENT, M/M_{yp}				
		0.215	1.074	1.289	1.397	1.289
Nelder-Mead's	(m)	360.82*	-	-	-	-
Simplex	(e)	18.73	-	-	-	-
Method	(t)	13.63	-	-	-	-
Powell's Conjugate	(m)	1.00	3.00	1.00	1.00	2.44
Direction	(e)	1.83	10.59	2.16	1.74	5.75
Method	(t)	1.44	4.25	1.86	1.70	4.97
Fletcher-Reeves' method	(m)	2.18	6.25	10.31*	-	-
with restarts after the	(e)	1.22	7.06	9.54	-	-
first N+2 iterations only	(t)	1.00	2.99	10.84	-	-
Method of	(m)	272.73*	-	-	-	-
Steepest Descent	(e)	146.95	-	-	-	-
	(t)	89.89	-	-	-	-
Davidon's Variance	(m)	8.18	26.25	14.06*	-	-
Method with	(e)	3.51	22.57	9.72	-	-
$\underline{H}_0 = \kappa \underline{I}$ where $\kappa=0.1$	(t)	2.40	8.53	8.86	-	-
Fletcher's Method	(m)	9.45	1.75	1.17	1.44	2.67
with $\underline{H}_0=\underline{I}$ only	(e)	4.28	2.13	1.00	1.00	3.09
initially	(t)	3.20	1.34	1.00	1.00	2.82
Davidon-Fletcher-Powell's	(m)	1.55	1.00	2.29	3.52	1.00
Method with $\underline{H}_0=\underline{I}$ only	(e)	1.00	1.00	2.09	2.67	1.00
initially	(t)	1.02	1.00	2.49	2.65	1.00

*Failed to Converge for given load increment

(m)...Weighting factor for the number of minimizations.
(e)...Weighting factor for the number of equivalent member energy evaluations.
(t)...Weighting factor for the elapsed CPU time.

Table 4.- Performance Data for an Elastic-Plastic Cantilever Beam Subjected to a Tip Bending Moment with Moderately Large Rotations Using Central Difference Gradients.

MINIMIZATION ALGORITHM		APPLIED TIP BENDING MOMENT, M/M_{yp}				
		0.215	1.074	1,289	1.397	1.289
Powell's Conjugate Direction Method	(m)	1.00	2,48**	-	-	-
	(e)	1,18	5,62	-	-	-
	(t)	1.10	5.14	-	-	-
Fletcher-Reeves' method with restarts after the first N+2 iterations only	(m)	4.91	14.33**	-	-	-
	(e)	3.03	16.21	-	-	-
	(t)	2.75	14,89	-	-	-
Davidon's Variance Method with $\underline{H}_0=\kappa\underline{I}$ where $\kappa=0.1$	(m)	17.24*	-	-	-	-
	(e)	7.21	-	-	-	-
	(t)	6,78	-	-	-	-
Fletcher's Method with $\underline{H}_0=\underline{I}$ only initially	(m)	2.28	1.48	1.00	1.00	1.00
	(e)	1.00	1.35	1.00	1.00	1.00
	(t)	1.00	1.30	1.00	1.00	1.00
Davidon-Fletcher-Powell's Method with $\underline{H}_0=\underline{I}$ only initially	(m)	10.47	1.00	2.34*		
	(e)	6.05	1.00	2,83		
	(t)	6.02	1.00	3.12		

*Failed to converge for given load increment

**Converged but analysis terminated

(m)...Weighting factor for the number of minimizations.
(e)...Weighting factor for the number of equivalent member energy evaluations.
(t)...Weighting factor for the elapsed CPU time.

Table 5.- Performance Data for an Impulsively Loaded Elastic-Plastic Clamped Beam with Moderately Large Rotations Using Central Difference Gradients.

MINIMIZATION ALGORITHM		TIME (IN GROUPS OF TEN TIME STEPS; $t_0=0.0$, $\Delta t=10^{-6}$ SECONDS) I	II	III	IV
Nelder-Mead's Simplex Method	(m)	135.41*	-	-	-
	(e)	6.44	-	-	-
	(t)	2.44	-	-	-
Powell's Conjugate Direction Method	(m)	1.00	1.26†	-	-
	(e)	2.55	5.92	-	-
	(t)	1.77	4.05	-	-
Fletcher-Reeves' Method with restarts after the first N+2 iterations only	(m)	5.41	10.37†	-	-
	(e)	2.93	8.15	-	-
	(t)	2.13	7.10	-	-
Method of Steepest Descent	(m)	21.89†	-	-	-
	(e)	11.75	-	-	-
	(t)	7.12	-	-	-
Davidon's Variance Method with $\underline{H}_0=\kappa\underline{I}$ where $\kappa=0.1$	(m)	15.07**	-	-	-
	(e)	6.63	-	-	-
	(t)	4.15	-	-	-
Fletcher's Method with $\underline{H}_0=\underline{I}$ only initially	(m)	2.52	1.00	1.00	1.00
	(e)	1.53	1.00	1.00	1.00
	(t)	1.07	1.00	1.00	1.00
Davidon-Fletcher-Powell's Method with $\underline{H}_0=\underline{I}$ only initially	(m)	2.00	2.39	3.46	2.02
	(e)	1.00	2.08	3.26	1.96
	(t)	1.00	1.91	2.65	1.69

*Failed to converge (for given time step) at $t=6.0\times10^{-6}$

**Failed to converge (for given time step) at $t=9.0\times10^{-6}$

†Converged but analysis terminated

(m)...Weighting factor for the number of minimizations.
(e)...Weighting factor for the number of equivalent member energy evaluations.
(t)...Weighting factor for the elapsed CPU time.

Table 6. - Performance Data for the Elastica Problem Using Analytic Gradients

MINIMIZATION ALGORITHM		APPLIED AXIAL LOAD, P/P_{cr} 0.1	0.2	0.3	0.4
Fletcher-Reeves' Method with restarts after the first N+2 iterations only	(m)	2.27	16.82	15.06	22.54
	(e)	2.80	34.69	20.94	32.31
	(t)	1.99	17.96	12.13	16.23
Fletcher's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	2.97	1.00	1.41	2.15
	(e)	1.35	1.00	1.00	1.63
	(t)	1.36	1.00	1.00	1.30
Davidon-Fletcher-Powell's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	1.00	1.27	1.00	1.00
	(e)	1.00	1.81	1.22	1.00
	(t)	1.00	1.67	1.16	1.00

(m)... Weighting factor for the number of minimizations.
(e)... Weighting factor for the number of equivalent member energy evaluations.
(t)... Weighting factor for the elapsed CPU time.

Table 7.- Performance Data for an Elastic Cantilever Beam Subjected to a Tip Bending Moment with Moderately Large Rotations Using Analytic Gradients

MINIMIZATION ALGORITHM		APPLIED TIP BENDING MOMENT, M/M_{yp}	
		0.215	1.074
Fletcher's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	1.00	1.23
	(e)	1.00	1.00
	(t)	1.00	1.00
Davidon-Fletcher-Powell's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	3.66	1.00
	(e)	7.33	1.24
	(t)	6.64	1.27

(m)... Weighting factor for the number of minimizations.
(e)... Weighting factor for the number of equivalent member energy evaluations.
(t)... Weighting factor for the elapsed CPU time.

Table 8.- Performance Data for an Elastic-Plastic Cantilever Beam Subjected to a Tip Bending Moment with Small and Geometrically Linear Deformations and Rotations Using Analytic Gradients.

MINIMIZATION ALGORITHM		APPLIED TIP BENDING MOMENT, M/M_{yp}			
		0.215	1.074	1.289	1.397
Fletcher-Reeves' Method with restart after the first N+2 iterations only	(m)	2.83	19.00	150.20*	--
	(e)	2.12	21.14	139.81	--
	(t)	1.42	3.19	104.15	--
Fletcher's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	8.67	3.50	1.20	1.00
	(e)	3.43	2.29	3.79	4.75
	(t)	2.46	1.25	2.38	2.12
Davidon-Fletcher-Powell's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	1.00	1.00	1.00	2.00
	(e)	1.00	1.00	1.00	1.00
	(t)	1.00	1.00	1.00	1.00

*Failed to converge for given load increment.

(m)... Weighting factor for the number of minimizations.
(e)... Weighting factor for the number of equivalent member energy evaluations.
(t)... Weighting factor for the elapsed CPU time.

Table 9.- Ratio of the Computational Effort with Analytic Gradients to That of Central Difference Gradients for the Elastica Problem.

MINIMIZATION ALGORITHM		APPLIED AXIAL LOAD, P/P_{cr}			
		0.1	0.2	0.3	0.4
Fletcher-Reeves' Method with restarts after the first N+2 iterations only	(m)	0.919	0.688	0.952*	1.089*
	(e)	0.518	0.359	0.452	0.495
	(t)	0.467	0.322	0.404	0.456
Fletcher's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	1.000	1.000	1.000	1.000
	(e)	0.320	0.323	0.323	0.323
	(t)	0.347	0.419	0.417	0.384
Davidon-Fletcher-Powell's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	0.833	0.778	1.417	0.929
	(e)	0.424	0.306	0.736	0.367
	(t)	0.460	0.390	0.834	0.512

*Convergence was not achieved for the third load step using central difference gradients; therefore, the central difference results of the second load step were used in calculating this ratio.

(m)... Ratio of the number of minimizations.
(e)... Ratio of the number of equivalent member energy evaluations.
(t)... Ratio of the elpased CPU time.

Table 10: Ratio of the Computational Effort with Analytic Gradients to That of Central Difference Gradients for an Elastic Cantilever Beam Subjected to a Tip Bending Moment with Moderately Large Rotations.

MINIMIZATION ALGORITHM		APPLIED TIP BENDING MOMENT, M/M_{yp} 0.215	1.074
Fletcher's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	0.992	0.939
	(e)	0.345	0.323
	(t)	0.355	0.331
Davidon-Fletcher-Powell's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	0.789	1.136
	(e)	0.419	0.539
	(t)	0.391	0.546

(m)... Ratio of the number of minimizations.
(e)... Ratio of the number of equivalent member energy evaluations.
(t)... Ratio of the elapsed CPU time.

Table 11.- Ratio of the Computational Effort with Analytic Gradients to That of Central Difference Gradients for an Elastic-Plastic Cantilever Beam Subjected to a Tip Bending Moment with Small and Geometrically Linear Deformations and Rotations.

MINIMIZATION ALGORITHM		APPLIED TIP BENDING MOMENT, M/M_{yp} 0.215	1.074	1.289	1.397
Fletcher-Reeves' Method with restart after the first N+2 iterations only	(m)	1.417	1.520	--*	--
	(e)	0.741	0.776	--	--
	(t)	0.789	0.786	--	--
Fletcher's Method with $\underline{H}_o = \underline{I}$ initially	(m)	1.000	1.000	0.146	0.026
	(e)	0.342	0.278	0.239	0.144
	(t)	0.428	0.686	0.164	0.102
Davidon-Fletcher-Powell's Method with $\underline{H}_o = \underline{I}$ only initially	(m)	0.706	0.500	0.063	0.021
	(e)	0.425	0.259	0.030	0.011
	(t)	0.548	0.737	0.028	0.018

*Failed to converge using both analytic and central difference gradients for given load increment.

(m)... Ratio of the number of minimizations.
(e)... Ratio of the number of equivalent member energy evaluations.
(t)... Ratio of the elapsed CPU time.

7. APPENDIX

UNCONSTRAINED MINIMIZATION ALGORITHMS

A.1 Introduction

The majority of the unconstrained minimization algorithms considered in this study hypothesize that the function to be minimized has a quadratic form. The rationale of this hypothesis is revealed by examining the Taylor series expansion of a general function of N variables $f(\underline{x})$ about its minimum point $\underline{x}_m$. That is,

$$f(\underline{x}) \sim f(\underline{x}_m) + \sum_{k=1}^{N} \left[\frac{\partial f}{\partial x_k}\right]_{\underline{x}_m} (x_k - x_{mk})$$

$$+ \frac{1}{2!} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\frac{\partial^2 f}{\partial x_i \ \partial x_j}\right]_{\underline{x}_m} (x_i - x_{mi}) (x_j - x_{mj})$$

$$+ \text{ higher order terms} \qquad \text{(A.1)}$$

or in vector form

$$f(\underline{x}) \quad a + \underline{b}^T(\underline{x} - \underline{x}_m) + \frac{1}{2!} \ (\underline{x} - \underline{x}_m) \ \underline{A}(\underline{x} - \underline{x}_m)$$

$$+ \text{ higher order terms} \qquad \text{(A.2)}$$

Since at the minimum point $\underline{x}_m$, the gradient vanishes and the term $(\underline{x} - \underline{x}_m)$ approaches zero, the higher order terms become relatively insignificant thereby causing the Taylor series expansion of $f(\underline{x})$ to be dominated by the constant, linear and quadratic terms in the proximity of the minimum. This suggests that in the neighborhood of the minimum, many general functions can be approximately represented by a quadratic. The function to be minimized in this study is Eq. (3).

A prevalent feature of the techniques considered is the replacement of an N-dimensional minimization problem with a sequence of one-dimensional or linear minimizations along the line.

$$\underline{x}_{k+1} = \underline{x}_k + \alpha \, \underline{d}_k \tag{A.3}$$

such that $f(\underline{x}_{k+1})$ becomes a function of α alone when the variables $\underline{x}_k$ and the search direction $\underline{d}_k$ are both known.

The method utilized for the linear minimization depends on the computational sophistication of the particular algorithm invoked. If gradient information is available, Davidon's cubic interpolation method [26, 27, 28] is employed; if not, Powell's quadratic interpolation method [13, 27, 28] is engaged. Powell's method finds the minimum of a quadratic which is described by the values of the function at three points on the line specified by Eq. (A.3). On the other hand, Davidon's method minimizes a cubic which is defined by the values of the function and its directional derivative at two points along the line given by Eq. (A.3). Thus, Davidon's formulae are more elaborate and time consuming than those of Powell but Davidon's method ordinarily locates the minimum along the line in fewer iterations [27].

Each of the techniques previously listed will now be characterized with the steps of their algorithms being omitted. The details of the algorithms are available in the appropriate literature cited.

A.2 Nelder-Mead's Simplex Method

In using the simplex method to minimize a function of N variables, the function is first evaluated at N + 1 mutually equidistant points in the space of the N variables. These points are called the vertices of a regular simplex which in two dimensions corresponds to an equilateral triangle and in three dimensions to a regular tetrahedron. This direct search technique using regular simplices was devised by Himsworth, Spendley and Hext [25]. Nelder-Meld's simplex method [12] represents an extension of their work wherein the regularity of the simplex is abandoned. Also, additional flexibility for adaptation of the simplex to the local geometry of the surface is incorporated. The basic operations of the simplex method are reflection, expansion and contraction. The vertex of the simplex with the highest function value is replaced by another point obtained through a reflection of that vertex. Depending on the value of the function at this new vertex, another reflection may be required or the simplex may expand or contract. These operations are repeated until the simplex essentially collapses on the minimum.

For each of these operations, a corresponding coefficient has been defined: a reflection coefficient α; a contraction coefficient β; and, an expansion coefficient γ. The values assigned to each of these coefficients are not fixed; however, Nelder and Mead's recommendation that the user set $\alpha = 1$, $\beta = 1/2$ and $\gamma = 2$ was incorporated in this study. These values correspond to a simple reflection, halving when in difficulty and doubling when a useful direction is located.

One undesireable feature of the simplex method is that it requires a significant amount of computer storage when N is large. This is true because the coordinates of N + 1 vertices are stored throughout the minimization process.

A.3 Powell's Conjugate Direction Method

Another direct search method is the univariate method which seeks the minimum of a function of N variables by changing one variable at a time. The method of conjugate directions [13] is essentially a variation of the univariate method wherein the properties of a quadratic function are exploited. Neglecting the higher order terms, Eq. (A.2) is a quadratic function with a Hessian matrix $\underline{A}$ (i.e. $\frac{\partial^2 f}{\partial \underline{x}^2} = \underline{A}$). Search directions $\underline{d}_1, \ldots\ldots, \underline{d}_N \neq 0$ with the property

$$\underline{d}_i^T \; \underline{A} \; \underline{d}_j = 0, \; i \neq j$$

are said to be conjugate or orthogonal with respect to the weighting matrix $\underline{A}$. Therefore, if these direction vectors are generated such that

$$\underline{d}_i = \underline{x}_i - \underline{x}_{i-1}$$

Then it can be shown [28, 29] that they remain $\underline{A}$-conjugate and a conjugate direction has been defined.

In this technique, each iteration begins with a linear search along N linearly independent directions $\underline{d}_1$, $\underline{d}_2$, ..., $\underline{d}_N$, starting at the best known approximation to the location of the minimum. After an initial univariate search along each of the coordinate directions, conjugate directions are generated by making each iteration define a new search direction $\underline{d}$ and choosing $\underline{d}_2$, $\underline{d}_3$, ..., $\underline{d}_N$, $\underline{d}$ as the linearly independent directions for the next iteration. Applying this procedure to a positive definite quadratic form, all search directions will be mutually conjugate after N linear minimizations and the exact minimum will have been found. However, Powell modified his basic procedure because on occasions it may choose nearly dependent directions; particularly when minimizing a function of more than five independent variables [27]. To overcome this, he allows a direction other than $\underline{d}_1$ to be discarded. Therefore the N search directions can be chosen so as to be always linearly independent, even though, in some cases, the same N directions are used for two successive linear minimizations. As a consequence of this modification, one of the mutually conjugate directions may be discarded, thereby requiring more than N linear minimizations to find the exact minimum of a quadratic.

A.4 Method of Steepest Descent

The method of steepest descent, first proposed by Cauchy [30], is perhaps the oldest and most well known gradient technique for function minimization. It is based on moving in the opposite direction of the gradient vector for minimization. No other vector can locally reduce the function value as much since the gradient vector perpendicularly cuts adjacent contours of the function. While simple and stable, this method often converges slowly and in a zig zag fashion.

This is because the direction of steepest descent and the direction to the minimum may be nearly perpendicular [31]. Fletcher [32] attributes this to the failure of the steepest descent theory to adequately represent functions with minima. He further asserts that "the only functions for which the steepest descent property holds along the whole direction of search and which still have a minimum are those with spherical contours, and this does not adequately represent the minimum of a general function" [32]. Research to strengthen this method for eccentric functions or functions whose contours are distorted hyperspheres has been carried out [33]. However, as Fox [28] points out, these modifications are of little value when compared to more recent techniques for function minimization.

A.5 Fletcher-Reeves' Conjugate Gradient Method

The conjugate gradient method is a special case of the more general method of conjugate directions and was first developed by Hestenes and Stiefel [34]. In the conjugate gradient method, the search directions $\underline{d}_1, \ldots, \underline{d}_N$ are generated such that $\underline{d}_{i+1}$ represents a linear combination of the present gradient vector $\underline{g}_{i+1}$ and previous direction vectors $\underline{d}_1, \ldots, \underline{d}_i$. That is,

$$\underline{d}_{i+1} = -\underline{g}_{i+1} + \frac{\underline{g}^2_{i+1}\,\underline{d}_i}{\underline{g}_i^{\,2}}$$

With this scheme for defining new search directions, the property of these directions being $\underline{A}$-conjugate is retained.

The Fletcher-Reeves algorithm [14] converts the N-dimensional problem into a sequence of one-dimensional problems; however, the accuracy of the one-dimensional minimization is not as crucial for this method as it is for the Davidon-Fletcher-Powell algorithm [11]. For quadratic functions, Myers [29] has shown that the direction vectors generated by the Fletcher-Reeves method are the same theoretically as those generated by the Davidon-Fletcher-Powell method, provided that the initial step for each method is taken in the direction of the steepest descent. This method reduces the convergence difficulties encountered by the method of steepest descent. Thus, accelerated convergence can be attained by reinitializing the search directions to those of the steepest descent method after every N + 2 iterations. This becomes especially necessary for nonquadratic functions because the conjugate gradient technique tends to generate nearly dependent search directions after N or so iterations [11]. With regard to reinitialization, Fox proposes that if the function is not both highly eccentric and twisted and "if the starting point is known to be a poor approximation to the minimum of a nonquadratic function, it may pay to restart after the first N or so cycles should they fail to produce a minimum, but not restart thereafter" [28].

A.6 Jacobson-Oksman's Method

The Jacobson-Oksman method [15] of function minimization is an innovative and relatively new approach which is based not on quadratic functions of the type

$$f(\underline{x}) = \frac{1}{2}(\underline{x} - \underline{x}_m)^T \underline{A}(\underline{x} - \underline{x}_m) + \bar{\omega}$$

but on homogeneous functions of the form

$$f(\underline{x}) = \frac{1}{\gamma}(\underline{x} - \underline{x}_m)^T \; \underline{g}(\underline{x}) + \bar{\omega}$$

where

$\underline{A}$ is an N x N constant positive definite matrix

$\underline{x}_m$ is the location of the minimum

$\bar{\omega}$ is the minimum function value, $f(\underline{x}_m)$

γ is the degree of homogeneity.

Based on homogeneous functions, this algorithm is not concerned with either the Hessian matrix or approximations to its inverse. The method converges in N + 2 steps for homogeneous functions. However, only descent and not convergence has been proven for general functions. Therefore, the algorithm cannot be expected to converge in N + 2 steps for general functions. It is noteworthy to underscore the arduous task of programming the algorithm even though it requires neither the Hessian matrix nor a one-dimensional minimization scheme except to guarantee descent or stability.

A.7 Davidon-Fletcher-Powell's Variable Metric Method

Based on the original work of Davidon [26], Fletcher and Powell modified the variable metric method in order to exploit its quadratic convergence properties and its stability [17]. The essence of the method is in the formation of a sequence of positive definite matrices which are used in the determination of the search directions. Any positive definite matrix may be used to start the iteration. This matrix is then updated at each iteration and converges to the inverse of the Hessian matrix at the minimum. The variable metric method also replaces the N-dimensional minimization problem with a sequence of one-dimensional minimizations using Davidon's cubic interpolation method. Since gradient information is required for the variable metric method, Stewart [18] has developed another modification allowing gradients to be calculated by the finite difference technique if analytic gradients are unavailable. Based on the fact that an approximation to the inverse of the Hessian matrix is available, Stewart's modification extracts an approximation to the Hessian matrix which is used to compute the step size that will produce maximum accuracy in the finite difference gradient.

A.8 Davidon's Variance Method

Davidon's variance method [16], occasionally referred to as Davidon's second method, is a similar yet simpler minimization algorithm than his first method, the variable metric method [26]. Within this context, the term variance has been generalized to mean the inverse of the Hessian matrix of any function, which will be computed by successive estimates as in the original variable metric algorithm. Therefore, the problem of inverting the Hessian matrix is avoided. Also, the one-dimensional minimization subproblem is no longer required provided a decrease in the function is realized at least once every N iterations [11]. However, the algorithm's principal drawback is its sensitivity to the user's selection of three parameters α, β and κ where $0 < \alpha < 1 < \beta$ and $\kappa > 0$ [11, 16]. These parameters appear to be very much problem dependent and thereby limit the usefulness of this technique.

A.9 Fletcher's New Variable Metric Method

Derived from Davidon-Fletcher-Powell's variable metric algorithm, Fletcher's new variable metric method [19] dispenses with the linear search which is crucial to the Davidon-Fletcher-Powell algorithm and próvides a new updating formula for the approximation to the inverse of the Hessian matrix. Having discarded the linear search, the property of quadratic termination cannot be proven and is replaced by a property for quadratic functions requiring that the eigenvalues of the inverse of the Hessian approximation tend monotonically toward those of the inverse of the Hessian matrix. Also required at each iteration is a sufficiently large reduction in the function value in order to guarantee ultimate convergence. An addition to the algorithm is the retention of the linear search capability which is invoked when it is necessary for other reasons to use more than one evaluation of the function and gradient per iteration.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-2991		

4. Title and Subtitle	5. Report Date
Efficiency of Unconstrained Minimization Techniques in Nonlinear Analysis	May 1978
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Manohar P. Kamat and Norman F. Knight, Jr.	
	10. Work Unit No.
9. Performing Organization Name and Address	11. Contract or Grant No.
Virginia Polytechnic Institute and State University Department of Engineering Science and Mechanics Blacksburg, Virginia 24061	NGR 47-004-114
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
National Aeronautics and Space Administration Langley Research Center Hampton, Virginia 23665	Contractor Report
	14 Sponsoring Agency Code

15. Supplementary Notes

Langley technical monitor: Robert J. Hayduk

Topical report.

16. Abstract

The most frequently used unconstrained minimization algorithms have been critically evaluated for their effectiveness in solving structural problems involving geometric and material nonlinearities. The algorithms have been categorized as being zeroth, first, or second order depending upon the highest derivative of the function required by the algorithm. The sensitivity of these algorithms to the accuracy of derivatives clearly suggests using analytically derived gradients instead of finite difference approximations. The use of analytic gradients results in better control of the number of minimizations required for convergence to the exact solution.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Unconstrained Minimization Techniques Finite Element Nonlinear Analysis	Unclassified - Unlimited Subject Category 39

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	48	$4.50

*For sale by the National Technical Information Service, Springfield, Virginia 22161

NASA-Langley, 1978