PROCEEDINGS
FIFTH HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES

Edited By

Art Lew

Conference Held

January 11, 12, 13, 1972

University of Hawaii

Honolulu, Hawaii

Sponsored By:

Information Sciences Program and
Department of Electrical Engineering
University of Hawaii

Supported By:

U.S. Army Research Office, Durham
Regional Medical Program of Hawaii
U.S. Office of Naval Research
Digital Equipment Corporation
National Science Foundation

In Cooperation With:

The IEEE Groups on
Circuit Theory
Information Theory
Systems, Man and Cybernetics
The IEEE Computer Society
The IEEE Control Society
The Hawaii Section of IEEE

A NEW APPROACH TO NONLINEAR PROGRAMMING

John W. Bandler and C. Charalambous
Communications Research Laboratory and
Department of Electrical Engineering
McMaster University
Hamilton, Ontario, Canada

Abstract

A new approach to nonlinear programming is presented. The original nonlinear programming problem is formulated as an unconstrained minimax problem. Under reasonable restrictions it is shown that a point satisfying the necessary conditions for a minimax optimum also satisfies the Kuhn-Tucker necessary conditions for the original problem. A least pth type of objective function for minimization with extremely large values of p is proposed to solve the problem.

1. INTRODUCTION

A number of examples can be cited[1] when general nonlinear minimax approximation problems involving a finite point set have been reformulated as nonlinear programs and solved by well-established methods such as the penalty function method of Fiacco and McCormick[2,3]. Other methods of solving the resulting nonlinear programs include the repeated application of linear programming to suitably linearized versions of the nonlinear problem[4]. In the present paper, on the other hand, we show how conventional nonlinear programming problems can be formulated for solution as minimax problems, with several attendant advantages. To our knowledge, the scheme we have adopted does not appear to have been previously attempted.

2. THE NEW APPROACH

2.1 THE EQUIVALENT MINIMAX PROBLEM

The nonlinear programming problem can be stated as

$$\text{minimize } U(\underset{\sim}{\phi}) \tag{1}$$

subject to

$$g_i(\underset{\sim}{\phi}) \geq 0 \qquad i=1,2,\ldots,m \tag{2}$$

where U is the generally nonlinear objective function of k parameters

$$\underset{\sim}{\phi} \triangleq [\phi_1 \quad \phi_2 \quad \ldots \quad \phi_k]^T \tag{3}$$

and $g_1(\underset{\sim}{\phi})$, $g_2(\underset{\sim}{\phi})$, ..., $g_m(\underset{\sim}{\phi})$ are, in general, nonlinear functions of the parameters. We will assume that all the functions are continuous with continuous partial derivatives, and that the inequality constraints $g_i(\underset{\sim}{\phi}) \geq 0$, $i=1,2,\ldots,m$ satisfy the constraint qualification[5].

Consider the problem of minimizing the unconstrained function

$$V(\underset{\sim}{\phi},\underset{\sim}{\alpha}) = \max_{1\leq i\leq m} [U(\underset{\sim}{\phi}),\ U(\underset{\sim}{\phi}) - \alpha_i g_i(\underset{\sim}{\phi})] \tag{4}$$

where

$$\underset{\sim}{\alpha} \triangleq [\alpha_1 \quad \alpha_2 \quad \ldots \quad \alpha_m]^T \tag{5}$$

and

$$\alpha_i > 0 \qquad i=1,2,\ldots,m \tag{6}$$

Theorem. Sufficiently large $\alpha_1, \alpha_2, \ldots, \alpha_m$ can be found such that at the point $\underset{\sim}{\phi}^o$ where the necessary conditions for optimality of $V(\underset{\sim}{\phi},\underset{\sim}{\alpha})$ with respect to $\underset{\sim}{\phi}$ are satisfied the Kuhn-Tucker necessary conditions for optimality of the original nonlinear programming problem are also satisfied.

Proof. It can be shown[6] that the necessary conditions for optimality of $V(\phi,\alpha)$ for fixed α at some point ϕ^o, for

$$\nabla \triangleq [\frac{\partial}{\partial\phi_1} \; \frac{\partial}{\partial\phi_2} \; \cdots \; \frac{\partial}{\partial\phi_k}]^T \qquad (7)$$

yield

$$v_o \nabla U(\phi^o) + \sum_{i\in M} v_i (\nabla U(\phi^o) - \alpha_i \nabla g_i(\phi^o)) = 0 \qquad (8)$$

$$v_o + \sum_{i\in M} v_i = 1 \qquad (9)$$

$$v_i \geq 0 \qquad i \in M \qquad (10)$$

$$v_i = 0 \qquad i \notin M \qquad (11)$$

and, for

$$v_o > 0 \qquad (12)$$

also that

$$U(\phi^o) = U(\phi^o) - \alpha_i g_i(\phi^o) \qquad i \in M \qquad (13)$$

$$U(\phi^o) > U(\phi^o) - \alpha_i g_i(\phi^o) \qquad i \notin M \qquad (14)$$

Thus, the index set M, a subset of $\{1,2,\ldots,m\}$, denotes functions $g_i(\phi^o)$ belonging to the equal maxima of $U(\phi^o)$, $U(\phi^o) - \alpha_i g_i(\phi^o)$. Using (6), we have

$$g_i(\phi^o) = 0 \qquad i \in M \qquad (15)$$

$$g_i(\phi^o) > 0 \qquad i \notin M \qquad (16)$$

The conditions can thus be rewritten into the form

$$(v_o + \sum_{i=1}^{m} v_i)\nabla U(\phi^o) - \sum_{i=1}^{m} v_i \alpha_i \nabla g_i(\phi^o) = 0 \qquad (17)$$

$$\left.\begin{aligned} v_i\alpha_i g_i(\phi^o) &= 0 \qquad (18)\\ v_i\alpha_i &\geq 0 \qquad (19)\end{aligned}\right\} \quad i=1,2,\ldots,m$$

or

$$\nabla U(\phi^o) = \sum_{i=1}^{m} u_i \nabla g_i(\phi^o) \qquad (20)$$

$$\left.\begin{aligned} u_i g_i(\phi^o) &= 0 \qquad (21)\\ u_i &\geq 0 \qquad (22)\end{aligned}\right\} \quad i=1,2,\ldots,m$$

where

$$u_i \triangleq v_i \, \alpha_i \qquad i=1,2,\ldots,m \qquad (23)$$

Since $\alpha_i > 0$ and $v_i < 1$ for all $i=1,2,\ldots,m$ it follows that

$$u_i < \alpha_i \qquad i=1,2,\ldots,m \qquad (24)$$

but, from (9) and (23)

$$v_o + \sum_{i=1}^{m} \frac{u_i}{\alpha_i} = 1$$

or

$$\frac{1}{1-v_o} \sum_{i=1}^{m} \frac{u_i}{\alpha_i} = 1 \qquad (25)$$

Now, the relations (20) to (22) are the Kuhn-Tucker necessary conditions for optimality of the original nonlinear programming problem. The $u_1, u_2, \ldots, u_m$ are specific nonnegative numbers, so that sufficiently large positive $\alpha_1, \alpha_2, \ldots, \alpha_m$ must be chosen to provide $v_o>0$ satisfying (25). Clearly, some flexibility in their choice exists, but since $u_1, u_2, \ldots, u_m$ are not usually known in advance one may not be able to forecast their values. Threshold values can be found from

$$\sum_{i=1}^{m} \frac{u_i}{\alpha_i} < 1 \qquad (26)$$

It should be noted that if insufficiently large values of $\alpha_1, \alpha_2, \ldots, \alpha_m$ are chosen it can be shown that although a valid minimum of $V(\phi,\alpha)$ may be found, the constraints $g_i(\phi) \geq 0$ for all $i=1,2,\ldots,m$ may not be satisfied at that point.

2.2 POSSIBLE IMPLEMENTATION

One possible approach for minimizing $V(\phi,\alpha)$ with respect to ϕ and which the authors have used with some success is to assume

$$W(\phi,\alpha,\beta) = \max_{1\leq i\leq m} [U(\phi)+\beta, \; U(\phi)+\beta-\alpha_i g_i(\phi)]$$

$$= \lim_{p\to\infty} X(\phi,\alpha,\beta,p) \qquad (27)$$

where

$$X(\phi,\alpha,\beta,p) = ([w_o(U(\phi)+\beta)]^p + \sum_{i=1}^{m} [w_i(U(\phi)+\beta-\alpha_i g_i(\phi))]^p)^{\frac{1}{p}} \qquad (28)$$

and where

$$\beta \geq 0 \qquad (29)$$

$$w_o = \begin{cases} 0 & \text{for } U(\phi)+\beta < 0 \\ 1 & \text{for } U(\phi)+\beta \geq 0 \end{cases} \qquad (30)$$

$$w_i = \begin{cases} 0 & \text{for } U(\phi)+\beta - \alpha_i g_i(\phi) < 0 \\ 1 & \text{for } U(\phi)+\beta - \alpha_i g_i(\phi) \geq 0 \end{cases} \qquad (31)$$

$$w_o + \sum_{i=1}^{m} w_i \geq 1 \qquad (32)$$

$$p > 1 \qquad (33)$$

and proceed to minimize $X(\phi,\alpha,\beta,p)$ with respect to ϕ from an arbitrary starting point for selected α and β using a very large value of p.

In particular, it is noted that

$$V(\phi,\alpha) = W(\phi,\alpha,\beta) - \beta \quad (34)$$

The reason for β is to ensure that (32) is satisfied, i.e., that $X(\phi,\alpha,\beta,p) > 0$. If $X(\phi,\alpha,\beta,p)$ becomes 0, β may be increased, and the minimization procedure restarted.

If a minimum of $X(\phi,\alpha,\beta,p)$ with respect to ϕ is obtained for which some or all of the constraints are violated, the elements of α are increased, and the minimization procedure restarted.

3. CONCLUSIONS

A number of advantages are obtained by our approach. The first is that the minimization of V can be regarded as an essentially unconstrained problem and a number of simple and suitable methods are available for its solution. The second is that the starting point can, in principle, be anywhere. There is no need to distinguish between feasible and nonfeasible points. The third is that once suitable values for the α_i have been determined, one complete optimization yields the solution unlike, of course, penalty function methods. The fourth is that, as the authors have shown[7-9], minimax problems can be reformulated as least pth problems and easily and efficiently solved by using extremely large values of p in conjunction with gradient methods such as the Fletcher-Powell method. Finally, we note that nonlinear equality constraints can also be readily handled by our method.

4. ACKNOWLEDGEMENT

This work was supported by the National Research Council of Canada under grant A7239 and by a Frederick Gardner Cottrell grant from the Research Corporation.

5. REFERENCES

[1] A.D. Waren, L.S. Lasdon and D.F. Suchman, "Optimization in engineering design", Proc. IEEE, vol. 55, pp. 1885-1897, November 1967.

[2] A.V. Fiacco and G.P. McCormick, "The sequential unconstrained minimization technique for nonlinear programming, a primal-dual method", Management Science, vol. 10, pp. 360-366, January 1964.

[3] A.V. Fiacco and G.P. McCormick, "Computational algorithm for the sequential unconstrained minimization technique for nonlinear programming", Management Science, vol. 10, pp. 601-617, July 1964.

[4] M.R. Osborne and G.A. Watson, "An algorithm for minimax approximation in the nonlinear case", Computer J., vol. 12, pp. 63-68, February 1969.

[5] W.I. Zangwill, Nonlinear Programming. Englewood Cliffs, N.J.: Prentice-Hall, 1969, p. 39.

[6] J.W. Bandler, "Conditions for a minimax optimum", IEEE Trans. Circuit Theory, vol. CT-18, pp. 476-479, July 1971.

[7] J.W. Bandler and C. Charalambous, "Generalized least pth objectives for networks and systems", Proc. 14th Midwest Symposium on Circuit Theory, Denver, Colo., pp. 10.5.1-10.5.10, May 1971.

[8] J.W. Bandler and C. Charalambous, "On conditions for optimality in least pth approximation with $p\to\infty$", 9th Allerton Conf. on Circuit and System Theory, Urbana, Ill., October 1971.

[9] J.W. Bandler and C. Charalambous, "Practical least pth approximation with extremely large values of p", 5th Asilomar Conf. on Circuits and Systems, Pacific Grove, Calif., November 1971.