JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 13, No. 6, 1974

# Nonlinear Programming Using Minimax Techniques[1]

J. W. BANDLER[2] AND C. CHARALAMBOUS[3]

Communicated by G. L. Nemhauser

**Abstract.** A minimax 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 $p$th type of objective function for minimization with extremely large values of $p$ is proposed to solve the problem. Several numerical examples compare the present approach with the well-known SUMT method of Fiacco and McCormick. In both cases, a recent minimization algorithm by Fletcher is used.

**Key Words.** Nonlinear programming, minimax approximation, least-square methods, function minimization, penalty function methods.

## 1. Introduction

A number of examples can be cited (Ref. 1) when general nonlinear minimax approximation problems involving a finite point set have been reformulated as nonlinear programs and solved by well-established

[1] This paper is based on work presented at the 5th Hawaii International Conference on System Sciences, Honolulu, Hawaii, 1972. The authors are greatly indebted to V. K. Jha for his programming assistance and J. H. K. Chen who obtained some of the numerical results. This work was supported in part by the National Research Council of Canada under Grant No. A7239, by a Frederick Gardner Cottrell Grant from the Research Corporation, and through facilities and support from the Communications Research Laboratory of McMaster University.

[2] Associate Professor of Electrical Engineering, Department of Electrical Engineering, McMaster University, Hamilton, Ontario, Canada.

[3] Post-Doctoral Fellow, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada.

methods, such as the barrier function method of Fiacco and McCormick (Refs. 2 and 3). Other methods for solving the resulting nonlinear programs include the repeated application of linear programming to suitably linearized versions of the nonlinear problem (Ref. 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 particular scheme that we have adopted does not appear to have been attempted previously, although some exact penalty function methods for solving nonlinear programs have already appeared and been discussed in the literature (Refs. 5–8).

## 2. Present Approach

**2.1. Nonlinear Programming Problem.** The nonlinear programming problem can be stated as follows:

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

subject to

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

where $U$ is the generally nonlinear objective function of $k$ parameters $\phi$, where

$$\phi \triangleq \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_k \end{bmatrix}, \tag{3}$$

and $g_1(\phi)$, $g_2(\phi)$,..., $g_m(\phi)$ 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 $g_i(\phi) \geqslant 0$, $i = 1, 2,..., m$, are such that a Kuhn–Tucker solution exists (Ref. 9).

**2.2. Equivalent Minimax Problem.** Consider the problem of minimizing the unconstrained function

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

where

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

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

**Theorem 2.1.** If the Kuhn–Tucker necessary conditions for optimality of the nonlinear programming problem are satisfied at $\phi^0$, then positive $\alpha_1, \alpha_2, \ldots, \alpha_m$ can be found satisfying

$$\sum_{i=1}^{m} (u_i/\alpha_i) < 1 \tag{7}$$

such that $\phi^0$ satisfies the necessary conditions for optimality of $V(\phi, \alpha)$ with respect to $\phi$ (Ref. 10), where $u_1, u_2, \ldots, u_m$ are the Kuhn–Tucker multipliers.

**Proof.** The Kuhn–Tucker necessary conditions for optimality of the original nonlinear programming problem are

$$\nabla U(\phi^0) = \sum_{i=1}^{m} u_i \nabla g_i(\phi^0), \tag{8}$$

where

$$\nabla \triangleq [\partial/\partial\phi_1, \partial/\partial\phi_2, \ldots, \partial/\partial\phi_k]^T, \tag{9}$$

$$M \triangleq \{i \mid g_i(\phi^0) = 0, i = 1, 2, \ldots, m\}, \tag{10}$$

$$u_i \geqslant 0, \qquad i \in M, \tag{11}$$

$$u_i = 0, \qquad i \notin M, \tag{12}$$

and

$$g_i(\phi^0) = 0, \qquad i \in M, \tag{13}$$

$$g_i(\phi^0) > 0, \qquad i \notin M. \tag{14}$$

Assuming (7), we have

$$v_0 + \sum_{i=1}^{m} (u_i/\alpha_i) = 1, \tag{15}$$

where

$$v_0 > 0. \tag{16}$$

Let

$$v_i \triangleq u_i/\alpha_i, \qquad i = 1, 2, \ldots, m. \tag{17}$$

Then, (11), (12), and (15) become, respectively,

$$v_i \geqslant 0, \qquad i \in M, \tag{18}$$

$$v_i = 0, \qquad i \notin M, \tag{19}$$

$$v_0 + \sum_{i=1}^{m} v_i = 1. \tag{20}$$

Also using (8), (13), (14), and (15), we obtain,

$$v_0 \nabla U(\phi^0) + \sum_{i=1}^{m} v_i(\nabla U(\phi^0) - \alpha_i \nabla g_i(\phi^0)) = 0, \tag{21}$$

$$U(\phi^0) = U(\phi^0) - \alpha_i g_i(\phi^0), \qquad i \in M, \tag{22}$$

$$U(\phi^0) > U(\phi^0) - \alpha_i g_i(\phi^0), \qquad i \notin M. \tag{23}$$

Now, relations (18)–(23) are the necessary conditions for optimality of $V(\phi, \alpha)$ for fixed $\alpha$ at the point $\phi^0$ (Ref. 10). 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 satisfy inequality (7). Clearly, some flexibility in their choice exists, but since $u_1, u_2, \ldots, u_m$ are not known in advance one may not, in general, be able to forecast their values. 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 $g_i(\phi) \geqslant 0$ for all $i = 1, 2, \ldots, m$ may not be satisfied at that point.

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

$$W(\phi, \alpha, \beta) = \max_{1 \leqslant i \leqslant m} [U(\phi) + \beta, U(\phi) + \beta - \alpha_i g_i(\phi)] = \lim_{p \to \infty} X(\phi, \alpha, \beta, p), \tag{24}$$

where

$$X(\phi, \alpha, \beta, p) = ([w_0(U(\phi) + \beta)]^p + \sum_{i=1}^{m} [w_i(U(\phi) + \beta - \alpha_i g_i(\phi))]^p)^{1/p}, \tag{25}$$

and where

$$\beta \geqslant 0, \tag{26}$$

$$w_0 = \begin{cases} 0 & \text{for} \quad U(\phi) + \beta < 0, \\ 1 & \text{for} \quad U(\phi) + \beta \geqslant 0, \end{cases} \tag{27}$$

$$w_i = \begin{cases} 0 & \text{for} \quad U(\phi) + \beta - \alpha_i g_i(\phi) < 0, \\ 1 & \text{for} \quad U(\phi) + \beta - \alpha_i g_i(\phi) \geqslant 0, \end{cases} \tag{28}$$

$$w_0 + \sum_{i=1}^{m} w_i \geqslant 1, \tag{29}$$

$$p > 1, \tag{30}$$

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

In particular, it is noted that

$$V(\phi, \alpha) = W(\phi, \alpha, \beta) - \beta. \tag{31}$$

The reason for $\beta$ is to ensure that (29) is satisfied, i.e., that $X(\phi, \alpha, \beta, p) > 0$. If $X(\phi, \alpha, \beta, p)$ becomes 0, $\beta$ may be increased, and the minimization procedure restarted.

If a minimum of $X(\phi, \alpha, \beta, p)$ with respect $\phi$ is obtained for which some or all of the constraints are violated, the elements of $\alpha$ are increased, and the minimization procedure is restarted. In practice, a tolerance for violated constraints should be specified.

**2.4. Comments.** 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 $\alpha_i$ have been determined, one complete optimization yields the solution unlike, of course, conventional barrier function methods. More precisely, if the minimax problems are reformulated as least $p$th problems (Refs. 11–13), one complete optimization yields an approximate solution, depending on the magnitude of $p$ as in barrier function methods where the approximation depends on the size of the controlling parameter. Finally, we note that nonlinear equality constraints can also be readily handled by our method.

## 3. Examples

To illustrate some of the concepts discussed in this paper and to provide some means of assessing the usefulness of our approach to nonlinear programming, a number of numerical examples, which have already received attention in the literature on optimization, will now be considered. A CDC 6400 computer was used throughout. Fletcher's recent minimization algorithm (Ref. 14) is also employed throughout. One function evaluation includes evaluation of all first derivatives.

**3.1. Post Office Parcel Problem.** (Ref. 15). For the post office parcel problem,

$$U = -\phi_1\phi_2\phi_3,$$

subject to

$$\phi_1 \geqslant 0, \qquad \phi_2 \geqslant 0, \qquad \phi_3 \geqslant 0,$$

$$20 - \phi_1 \geqslant 0, \qquad 11 - \phi_2 \geqslant 0, \qquad 42 - \phi_3 \geqslant 0,$$

$$72 - \phi_1 - 2\phi_2 - 2\phi_3 \geqslant 0.$$

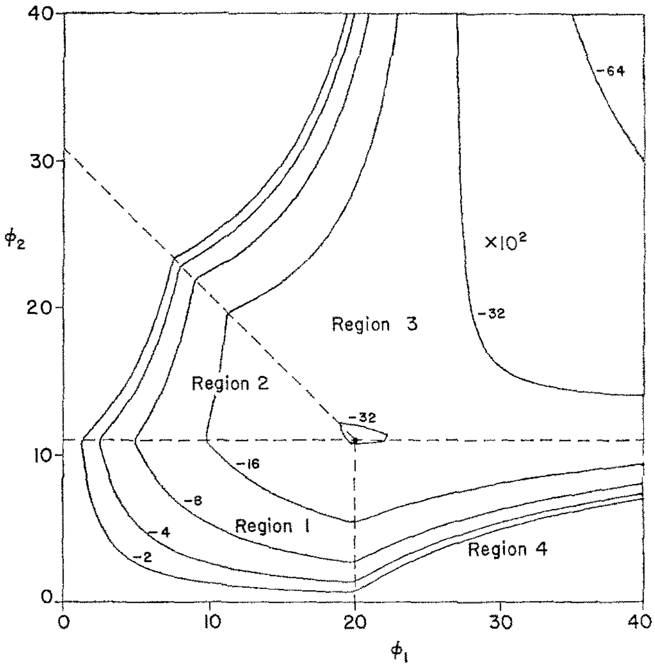


Fig. 1. Contours for post office parcel problem when $\alpha = 200$.

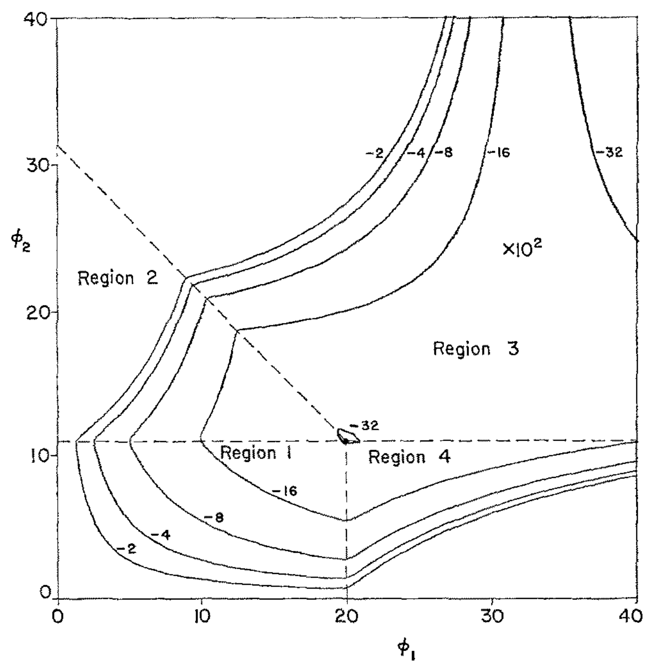


Fig. 2. Contours for post office parcel problem when $\alpha = 245$.

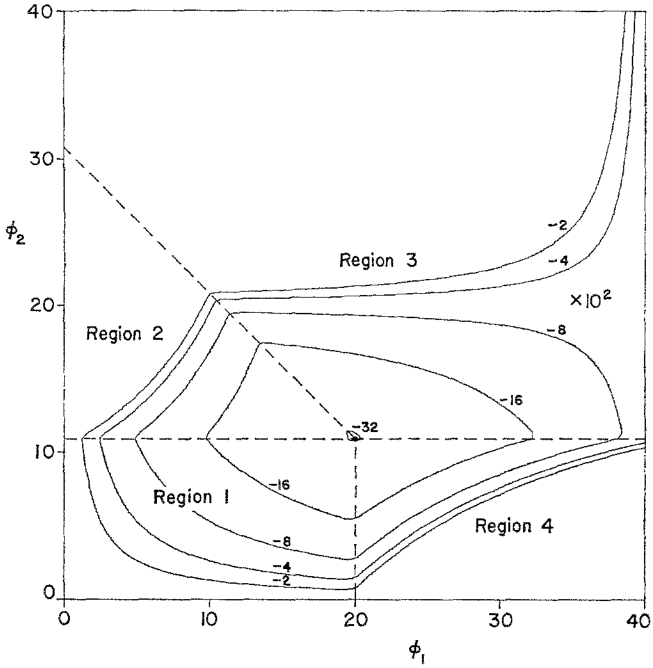


Fig. 3. Contours for post office parcel problem when $\alpha = 300$.

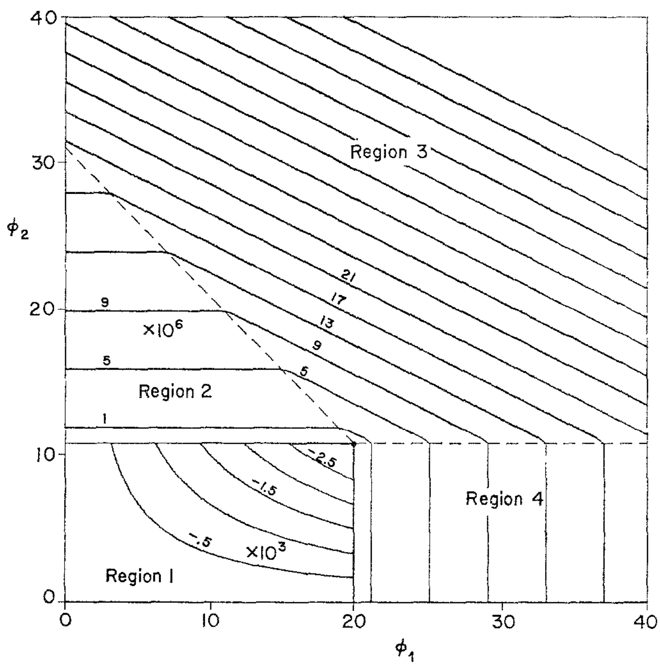


Fig. 4. Contours for post office parcel problem when $\alpha = 10^6$.

The solution is

$$U = -3300, \qquad \phi_1 = 20, \qquad \phi_2 = 11, \qquad \phi_3 = 15.$$

We confine ourselves on this problem to showing contours of $V$ against $\phi_1$ and $\phi_2$ for $\phi_3 = 15$ and for different sets of values of the $\alpha_i$. We let $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$. It can be shown that the threshold value of $\alpha$ is 245. Figure 1 shows the situation for $\alpha = 200$, Fig. 2 the situation for $\alpha = 245$, Fig. 3 the situation for $\alpha = 300$, and Fig. 4 the situation for $\alpha = 10^6$. In these figures we have in Region 1

$$V = U,$$

in Region 2

$$V = U - \alpha(11 - \phi_2),$$

in Region 3

$$V = U - \alpha(72 - \phi_1 - 2\phi_2 - 2\phi_3),$$

in Region 4

$$V = U - \alpha(20 - \phi_1),$$

and it is noted that only Region 1 is feasible. In Fig. 1 the minimum is nonfeasible. We note also that nonfeasible starting points are permitted although, as seen from Fig. 2, convergence to the desired minimum may not be guaranteed.

**3.2. Beale's Problem.** (Refs. 16 and 17). Here,

$$U = 9 - 8\phi_1 - 6\phi_2 - 4\phi_3 + 2\phi_1^2 + 2\phi_2^2 + \phi_3^2 + 2\phi_1\phi_2 + 2\phi_1\phi_3,$$

subject to

$$\phi_1 \geqslant 0, \qquad \phi_2 \geqslant 0, \qquad \phi_3 \geqslant 0,$$
$$3 - \phi_1 - \phi_2 - 2\phi_3 \geqslant 0,$$

The solution is

$$U = \frac{1}{9}, \qquad \phi_1 = \frac{4}{3}, \qquad \phi_2 = \frac{7}{9}, \qquad \phi_3 = \frac{4}{9}.$$

The SUMT method of Fiacco and McCormick (Refs. 2 and 3) was used to solve the problem by defining

$$B(\phi, r) = U(\phi) + r \sum_{i=1}^{m} [1/g_i(\phi)] \tag{32}$$

and minimizing $B$ with respect to $\phi$ for a strictly decreasing sequence of $r$ values. In Table 1, the same sequence as used by Kowalik, Osborne,

Table 1. Comparison on Beale problem for starting point $\phi_1 = \phi_2 = \phi_3 = 0.5$.

| Method | Fiacco–McCormick | Fiacco–McCormick | Bandler–Charalambous |
|---|---|---|---|
| $\phi_1$ | 1.33336 | 1.33336 | 1.333333 |
| $\phi_2$ | 0.77776 | 0.77776 | 0.777777 |
| $\phi_3$ | 0.44441 | 0.44441 | 0.444444 |
| $U$ | 0.11113 | 0.11113 | 0.11111 |
| Function evaluations | 127 | 67 | 48 |
| Parameter values | $r_1 = 1, r_{10} = 10^{-9}$ | $r = 10^{-9}$ | $\alpha = 1, \beta = 0, p = 10^5$ |

and Ryan (Ref. 17) was chosen. Table 2 shows results from a different starting point. From the nonfeasible point $\phi_1 = \phi_2 = \phi_3 = 1$, using the same $\alpha$, $\beta$, and $p$ as in Tables 1 and 2, our method reached $\phi_1 = 1.333330$, $\phi_2 = 0.7777813$, $\phi_3 = 0.4444436$, $U = 0.1111117$ in 42 function evaluations.

**3.3. The Rosen–Suzuki Problem.** (Refs. 16 and 17). In this case,

$$U = \phi_1^2 + \phi_2^2 + 2\phi_3^2 + \phi_4^2 - 5\phi_1 - 5\phi_2 - 21\phi_3 + 7\phi_4 ,$$

subject to

$$-\phi_1^2 - \phi_2^2 - \phi_3^2 - \phi_4^2 - \phi_1 + \phi_2 - \phi_3 + \phi_4 + 8 \geqslant 0,$$

$$-\phi_1^2 - 2\phi_2^2 - \phi_3^2 - 2\phi_4^2 + \phi_1 + \phi_4 + 10 \geqslant 0,$$

$$-2\phi_1^2 - \phi_2^2 - \phi_3^2 - 2\phi_1 + \phi_2 + \phi_4 + 5 \geqslant 0.$$

Table 2. Comparison on Beale problem for starting point $\phi_1 = \phi_2 = \phi_3 = 0.1$.

| Method | Fiacco–McCormick | Bandler–Charalambous |
|---|---|---|
| $\phi_1$ | 1.333336 | 1.333337 |
| $\phi_2$ | 0.7777763 | 0.7777868 |
| $\phi_3$ | 0.4444407 | 0.4444375 |
| $U$ | 0.1111126 | 0.1111115 |
| Function evaluations | 129 | 45 |
| Parameter values | $r_1 = 10^{-3}, r_9 = 10^{-11}$ | $\alpha = 1, \beta = 0, p = 10^5$ |

Table 3. Results for Rosen–Suzuki problem starting at $\phi_1 = \phi_2 = \phi_3 = \phi_4 = 0$ using $p = 10^5$.

| Parameter values | | Function evaluations | Objective function value |
|---|---|---|---|
| Initial | Final | | |
| $\alpha = 1, \beta = 100$ | $\alpha = 10, \beta = 100$ | 137 | $-43.9997$ |
| $\alpha = 10, \beta = 100$ | $\alpha = 10, \beta = 100$ | 78 | $-43.9958$ |
| $\alpha = 10, \beta = 1000$ | $\alpha = 10, \beta = 1000$ | 57 | $-43.9951$ |
| $\alpha = 1, \beta = 1000$ | $\alpha = 10, \beta = 1000$ | 95 | $-43.9950$ |
| $\alpha = 100, \beta = 1000$ | $\alpha = 100, \beta = 1000$ | 520 | $-43.9988$ |

The solution is

$$U = -44, \qquad \phi_1 = 0, \qquad \phi_2 = 1, \qquad \phi_3 = 2, \qquad \phi_4 = -1.$$

Table 3 shows the performance of our method for a number of different values of $\alpha$ and $\beta$. These parameters were increased by factors of 10, as necessary to obtain the desired solution. These results may be compared with the SUMT method from the same starting point, again generating the same sequence as used in Ref. 17. The value $U = -43.999$ was obtained in 145 function evaluations. Using $r = 10^{-9}$, the value $U = -42.173$ was obtained in 520 function evaluations, during which the Fletcher method had to be restarted three times.

**3.4. Quadratic Function With Equality Constraint.** This problem, which has been used by Fletcher (Ref. 6), has

$$U = \phi_1^2 + 4\phi_2^2,$$

subject to

$$\phi_1 + 2\phi_2 - 1 = 0.$$

The solution is

$$U = 0.5, \qquad \phi_1 = 0.5, \qquad \phi_2 = 0.25.$$

It can be shown that the threshold value of $\alpha$ is 1. Fig. 5 depicts contours of

$$V = \max[U, U + (\phi_1 + 2\phi_2 - 1), U - (\phi_1 + 2\phi_2 - 1)]$$

with respect to $\phi_1$ and $\phi_2$.

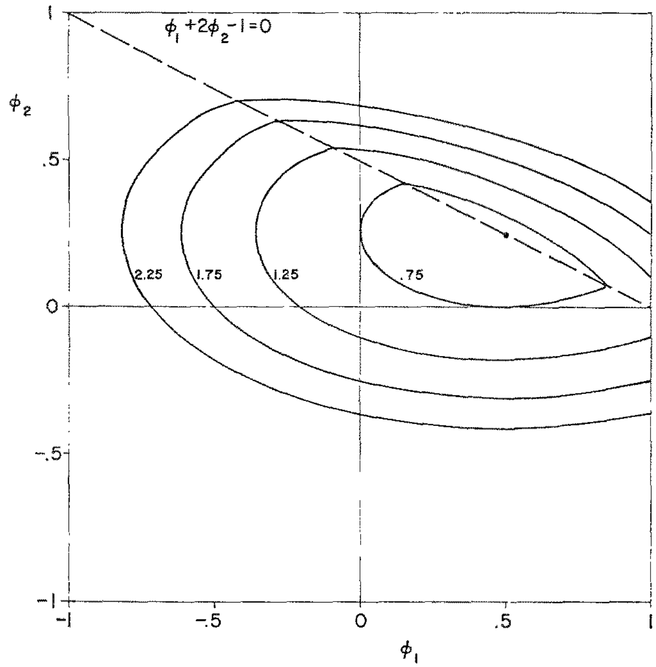


Fig. 5. Contours for quadratic function when $\alpha = 1$.

Using $\alpha = 10$, $\beta = 0$, and $p = 10^5$ and starting from $\phi_1 = \phi_2 = 1$, the result $U = 0.4999999$, $\phi_1 = 0.5000004$, $\phi_2 = 0.2499998$ was obtained in 37 function evaluations. For the SUMT method, we used

$$U(\phi) + (1/\sqrt{r})(\phi_1 + 2\phi_2 - 1)^2$$

beginning with $r = 10^{-3}$ and reducing $r$ by factors of 10 down to $10^{-10}$. The result $U = 0.4999950$, $\phi_1 = 0.4999975$, $\phi_2 = 0.2499987$ was obtained in 65 function evaluations. Starting with $r = 10^{-10}$, only four function evaluations yielded this solution.

## 4. Conclusions

As our numerical results indicate, the method of nonlinear programming that we have adopted is very promising. Crucial to efficiency is the method used to effectively minimize the unconstrained minimax objective function which we create and to rapidly determine sufficiently large $\alpha_i$ so that the desired feasible solution can be attained. Arbitrarily large values of $\alpha_i$ chosen initially may result in loss of efficiency due to poor scaling. Our use of Fletcher's minimization algorithm in conjunction with a value of $p$ as high as $10^5$ undoubtedly is a further obstacle to better results. Our experience on approximation problems indicates that values of $p$ of about 1000 yield highly acceptable results in a reasonable computing time. Our method does not depend so much on least $p$th approximation but strictly on available algorithms for solving nonlinear minimax approximation problems.

## References


1. WAREN, A. D., LASDON, L. S., and SUCHMAN, D. F., *Optimization in Engineering Design*, Proceedings of the IEEE, Vol. 55, No. 11, 1967.
2. FIACCO, A. V., and MCCORMICK, G. P., *The Sequential Unconstrained Minimization Technique for Nonlinear Programming, A Primal–Dual Method*, Management Science, Vol. 10, No. 2, 1964.
3. FIACCO, A. V., and MCCORMICK, G. P., *Computational Algorithm for the Sequential Unconstrained Minimization Technique for Nonlinear Programming*, Management Science, Vol. 10, No. 4, 1964.
4. OSBORNE, M. R., and WATSON, G. A., *An Algorithm for Minimax Approximation in the Nonlinear Case*, Computer Journal, Vol. 12, No. 1, 1969.
5. ZANGWILL, W. I., *Nonlinear Programming via Penalty Functions*, Management Science, Vol. 13, No. 5, 1967.
6. FLETCHER, R., *A Class of Methods for Nonlinear Programming with Termination and Convergence Properties*, Integer and Nonlinear Programming, Edited by J. Abadie, North-Holland Publishing Company, Amsterdam, Holland, 1970.
7. FLETCHER, R., *An Exact Penalty Function for Nonlinear Programming with Inequalities*, Mathematical Programming, Vol, 5, No. 2, 1973.
8. PIETRZYKOWSKI, T., *An Exact Potential Method for Constrained Maxima*, SIAM Journal of Numerical Analysis, Vol. 6, No. 2, 1969.
9. ZANGWILL, W. I., *Nonlinear Programming: A Unified Approach*, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.
10. BANDLER, J. W., *Conditions for a Minimax Optimum*, IEEE Transactions on Circuit Theory, Vol. CT-18, No. 4, 1971.
11. BANDLER, J. W., AND CHARALAMBOUS, C., *Practical Least pth Approximation with Extremely Large Values of p*, Conference Record of the Fifth Asilomar Conference on Circuits and Systems, Pacific Grove, California, 1971.
12. BANDLER, J. W., and CHARALAMBOUS, C., *Theory of Generalized Least pth Approximation*, IEEE Transactions on Circuit Theory, Vol. CT-19, No. 3, 1972.

13. Bandler, J. W., and Charalambous, C., *On Conditions for Optimality in Least pth Approximation with* $p \to \infty$, Journal of Optimization Theory and Applications, Vol. 11, No. 5, 1973.
14. Fletcher, R., *A New Approach to Variable Metric Algorithms*, Computer Journal, Vol. 13, No. 3, 1970.
15. Rosenbrock, H. H., *An Automatic Method for Finding the Greatest or Least Value of a Function*, Computer Journal, Vol. 3, No. 1, 1960.
16. Kowalik, J., and Osborne, M. R., *Methods for Unconstrained Optimization Problems*, American Elsevier Publishing Company, New York, New York, 1968.
17. Kowalik, J., Osborne, M. R., and Ryan, D. M., *A New Method for Constrained Optimization Problems*, Operations Research, Vol. 17, No. 6, 1969.