INTERNAL REPORTS IN

SIMULATION, OPTIMIZATION AND CONTROL

No. SOC-2

NONLINEAR PROGRAMMING USING MINIMAX TECHNIQUES

J.W. Bandler and C. Charalambous

June 1973

FACULTY OF ENGINEERING
McMASTER UNIVERSITY
HAMILTON, ONTARIO, CANADA

NONLINEAR PROGRAMMING USING MINIMAX TECHNIQUES

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

Abstract   A minimax approach to nonlinear programming is presented.

The original nonlinear programming problem is formulated as an un-

constrained minimax problem. Under reasonable restrictions it is

shown that a point satisfying the necessary conditions for a mini-

max 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. 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.

---

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

[3]Graduate Student, Department of Electrical Engineering, McMaster
University, Hamilton, Ontario, Canada.

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 methods such as the barrier function method of Fiacco and McCormick (Refs. 2 and 3). Other methods of 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 we have adopted does not appear to have been previously attempted, although some exact penalty function methods for solving nonlinear programs have already appeared and been discussed in the literature (Refs. 5 to 8).

2. The Present Approach

2.1 The Nonlinear Programming Problem

The nonlinear programming problem can be stated as

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

subject to

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

where U is the generally nonlinear objective function of k parameters ϕ,

where

$$\phi \overset{\Delta}{=} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_k \end{bmatrix} \tag{3}$$

and $g_1(\phi)$, $g_2(\phi)$, $\ldots$, $g_m(\phi)$ 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

$g_i(\phi) \geq 0$, $i=1,2,\ldots,m$ are such that a Kuhn-Tucker solution exists (Ref.9).

2.2 An Equivalent Minimax Problem

Consider the problem of minimizing the unconstrained function

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

where

$$\alpha \overset{\Delta}{=} [\alpha_1 \ \alpha_2 \ \ldots \ \alpha_m]^T \tag{5}$$

and

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

Theorem. If the Kuhn-Tucker necessary conditions for optimality of the nonlinear programming problem are satisfied at $\phi^o_{\sim}$, then positive $\alpha_1, \alpha_2, \ldots, \alpha_m$ can be found satisfying

$$\sum_{i=1}^{m} \frac{u_i}{\alpha_i} < 1 \tag{7}$$

such that $\phi^o_{\sim}$ satisfies the necessary conditions for optimality of $V(\phi_{\sim}, \alpha_{\sim})$ with respect to $\phi_{\sim}$, 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:

$$\nabla_{\sim} U(\phi^o_{\sim}) = \sum_{i=1}^{m} u_i \nabla_{\sim} g_i(\phi^o_{\sim}) \tag{8}$$

where,

$$\nabla_{\sim} \triangleq \begin{bmatrix} \dfrac{\partial}{\partial \phi_1} \\ \dfrac{\partial}{\partial \phi_2} \\ \vdots \\ \dfrac{\partial}{\partial \phi_k} \end{bmatrix} \tag{9}$$

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

$$u_i \geq 0 \qquad\qquad i \varepsilon M \tag{11}$$

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

and

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

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

Assuming (7),

$$v_o + \sum_{i=1}^{m} \frac{u_i}{\alpha_i} = 1 \tag{15}$$

where

$$v_o > 0 \tag{16}$$

Let

$$v_i \triangleq \frac{u_i}{\alpha_i} \quad i=1,2,\ldots,m \tag{17}$$

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

$$v_i \geq 0 \qquad i \varepsilon M \tag{18}$$

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

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

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

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

$$U(\underset{\sim}{\phi}^O) = U(\underset{\sim}{\phi}^O) - \alpha_i\, g_i(\underset{\sim}{\phi}^O) \qquad i\epsilon M \tag{22}$$

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

Now, relations (18) to (23) are the necessary conditions for

optimality of $V(\underset{\sim}{\phi},\underset{\sim}{\alpha})$ for fixed $\underset{\sim}{\alpha}$ at the point $\underset{\sim}{\phi}^O$. The $u_1, u_2, \ldots, u_m$ are

specific non-negative numbers, so that sufficiently large positive

$\alpha_1, \alpha_2, \ldots, \alpha_m$ must be chosen to satisfy inequality (7). Clearly some flex-

ibility 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 $V(\underset{\sim}{\phi},\underset{\sim}{\alpha})$ may be

found, the constraints $g_i(\phi) \geq 0$ for all $i=1,2,\ldots,m$ may not be satisfied

at that point.

2.3 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

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

$$= \lim_{p \to \infty} X(\phi,\alpha,\beta,p) \tag{24}$$

where

$$X(\phi,\alpha,\beta,p) = ([w_o(U(\phi) + \beta)]^p$$

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

and where

$$\beta \geq 0 \tag{26}$$

$$w_o = \begin{cases} 0 \text{ for } U(\phi)+\beta < 0 \\ 1 \text{ for } U(\phi)+\beta \geq 0 \end{cases} \tag{27}$$

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

$$w_o + \sum_{i=1}^{m} w_i \geq 1 \tag{29}$$

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

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

$$V(\phi,\alpha) = W(\phi,\alpha,\beta) - \beta \qquad (31)$$

The reason for β is to ensure that (29) 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. 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 α_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 pth problems (Ref. 11 to 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 The Post Office Parcel Problem (Ref. 15)

For the post office parcel problem

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

subject to

$$\phi_1 \geq 0, \quad \phi_2 \geq 0, \quad \phi_3 \geq 0$$

$$20 - \phi_1 \geq 0, \quad 11 - \phi_2 \geq 0, \quad 42 - \phi_3 \geq 0,$$

$$72 - \phi_1 - 2\phi_2 - 2\phi_3 \geq 0$$

The solution is

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

We confine ourselves on this problem to showing contours of V against ϕ_1 and ϕ_2 for $\phi_3 = 15$ and for different sets of values of the α_i. We let $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$. It can be shown that the threshold value of α is 245. Fig. 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 The Beale 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

$$\phi_1 \geq 0, \quad \phi_2 \geq 0, \quad \phi_3 \geq 0$$

$$3 - \phi_1 - \phi_2 - 2\phi_3 \geq 0$$

The solution is

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

The SUMT method of Fiacco and McCormick (Refs. 2 and 3) was

used to solve the problem by defining

$$B(\underset{\sim}{\phi}, r) = U(\underset{\sim}{\phi}) + r \sum_{i=1}^{m} \frac{1}{g_i(\underset{\sim}{\phi})} \tag{35}$$

and minimizing B with respect to ϕ for a strictly decreasing sequence of

r values.   In Table 1, the same sequence as used by Kowalik, Osborne 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 α, β

and p as in Tables 1 and 2, our method reached $\phi_1 = 1.333330$, $\phi_2 = 0.7777813$,

$\phi_3 = \cdot 4444436$, 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

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

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

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

The solution is

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

Table 3 shows the performance of our method for a number of

different values of α and β.   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 Constraints

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

$$U = 0.5, \quad \phi_1 = 0.5, \quad \phi_2 = 0.25$$

It can be shown that the threshold value of α 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 ϕ_1 and ϕ_2.

Using $\alpha = 10$, $\beta = 0$, and $p = 10^5$ and starting from $\phi_1 = \phi_2 = 1$,

the result U = 0.4999999, ϕ_1 = 0.5000004, ϕ_2 = 0.2499998 was obtained in

37 function evaluations. For the SUMT method we used

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

beginning with r = 10^{-3} and reducing r by factors of 10 down to 10^{-10}.

The result U = 0.4999950, ϕ_1 = 0.4999975, ϕ_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 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 α_i so that the desired feasible solution can be attained. Arbitrarily

large values of α_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 pth approximation but strictly on available algorithms for solving nonlinear minimax approximati 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. In Integer and Nonlinear Programming,

edited by J. Abadie, North-Holland Publishing Company, Amsterdam, 1970.

7. FLETCHER, R., <u>An Exact Penalty Function for Nonlinear Programming</u>

<u>with Inequalities</u>, Report TP 478, Atomic Energy Authority, Harwell,

England, 1972.

8. PIETRZYKOWSKI, T., <u>An Exact Potential Method for Constrained Maxima</u>,

SIAM Journal of Numerical Analysis, Vol. 6, No. 2, 1969.

9. ZANGWILL, W.I., <u>Nonlinear Programming: A Unified Approach</u>, Prentice-Hall,

Englewood Cliffs, N.J., 1969.

10. BANDLER, J.W., <u>Conditions for a Minimax Optimum</u>, IEEE Transactions on

Circuit Theory, Vol. CT-18, No. 4, 1971.

11. BANDLER, J.W. and CHARALAMBOUS, C., <u>Practical Least pth Approximation</u>

<u>with Extremely Large Values of p</u>, Conference Record of the Fifth

Asilomar Conference on Circuits and Systems, Pacific Grove, Calif, 1971.

12. BANDLER, J.W., and CHARALAMBOUS, C., <u>Theory of Generalized Least pth</u>

<u>Approximation</u>, 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→∞, Journal of Optimization Theory and Applications, to be published.

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, 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.

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

Methods	Fiacco-McCormick		Bandler-Charalambous
ϕ_1	1.33336		1.333333
ϕ_2	0.77776		0.777777
ϕ_3	0.44441		0.444444
U	0.11113		0.11111
Function Evaluations	127	67	48
Parameter Values	$r_1 = 1$ $r_{10} = 10^{-9}$	$r=10^{-9}$	$\alpha = 1$ $\beta = 0$ $p = 10^5$

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

Methods	Fiacco-McCormick	Bandler-Charalambous
ϕ_1	1.333336	1.333337
ϕ_2	0.7777763	0.7777868
ϕ_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

Figure Captions

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 $\alpha=10^6$.
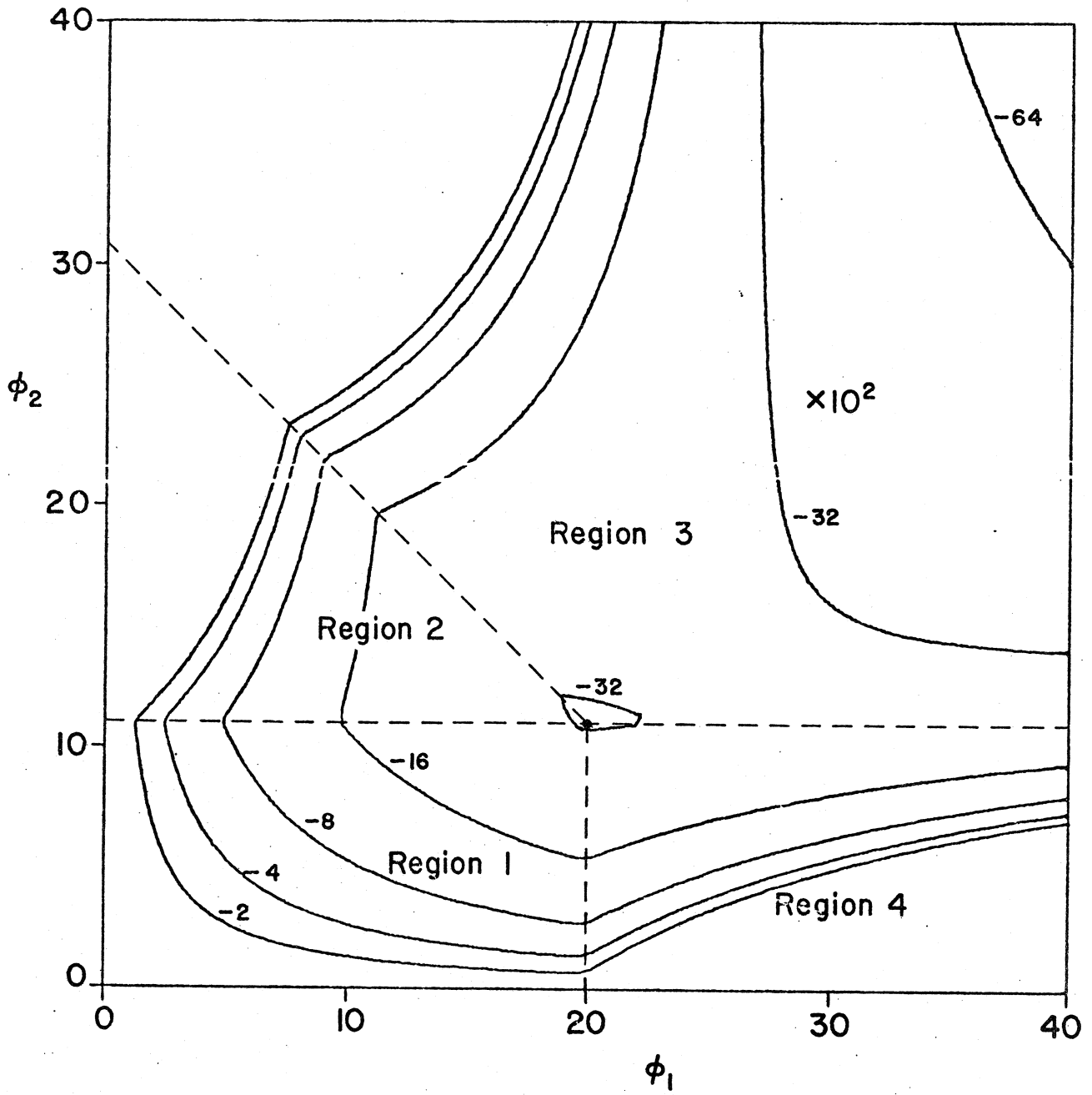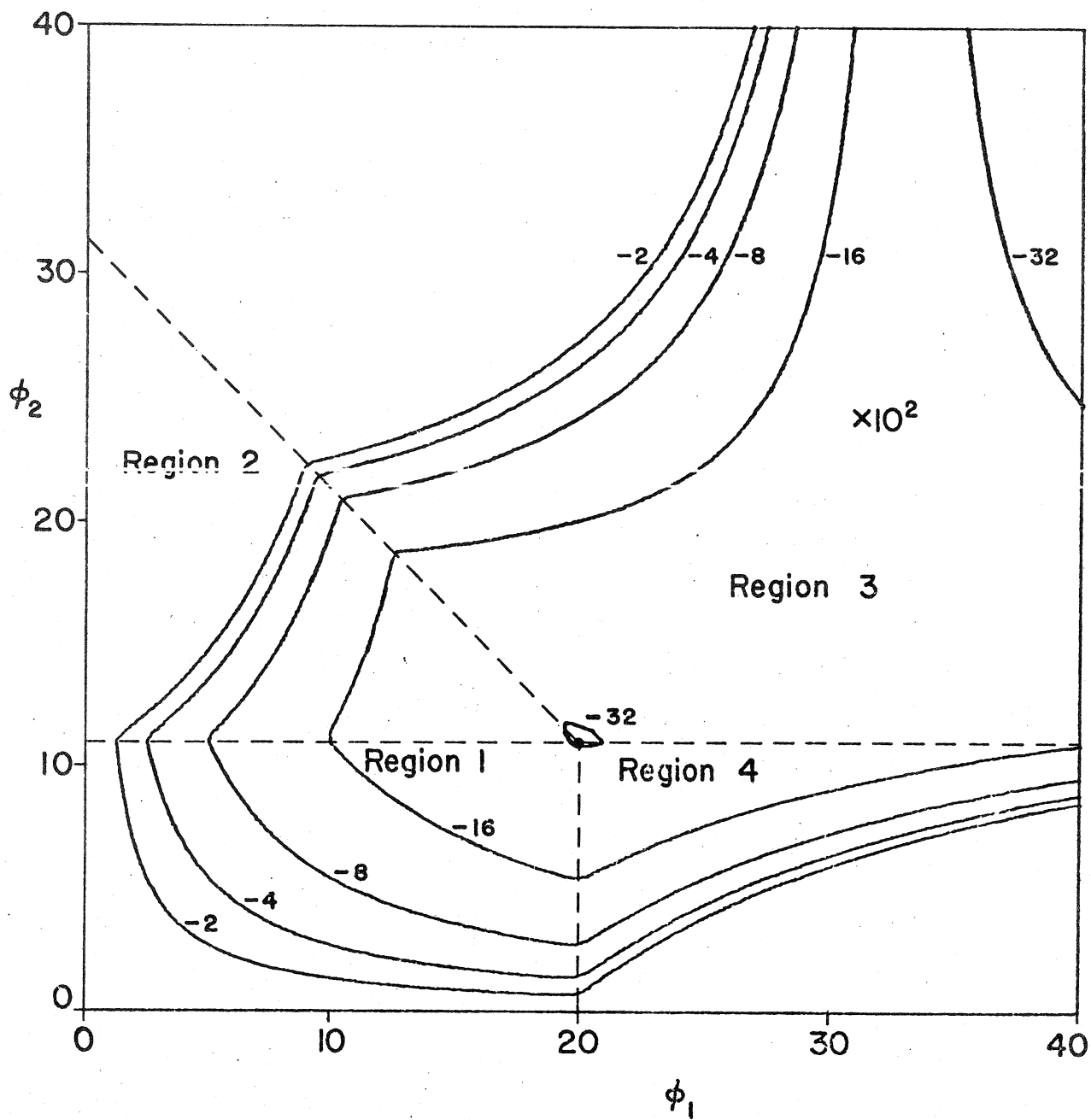
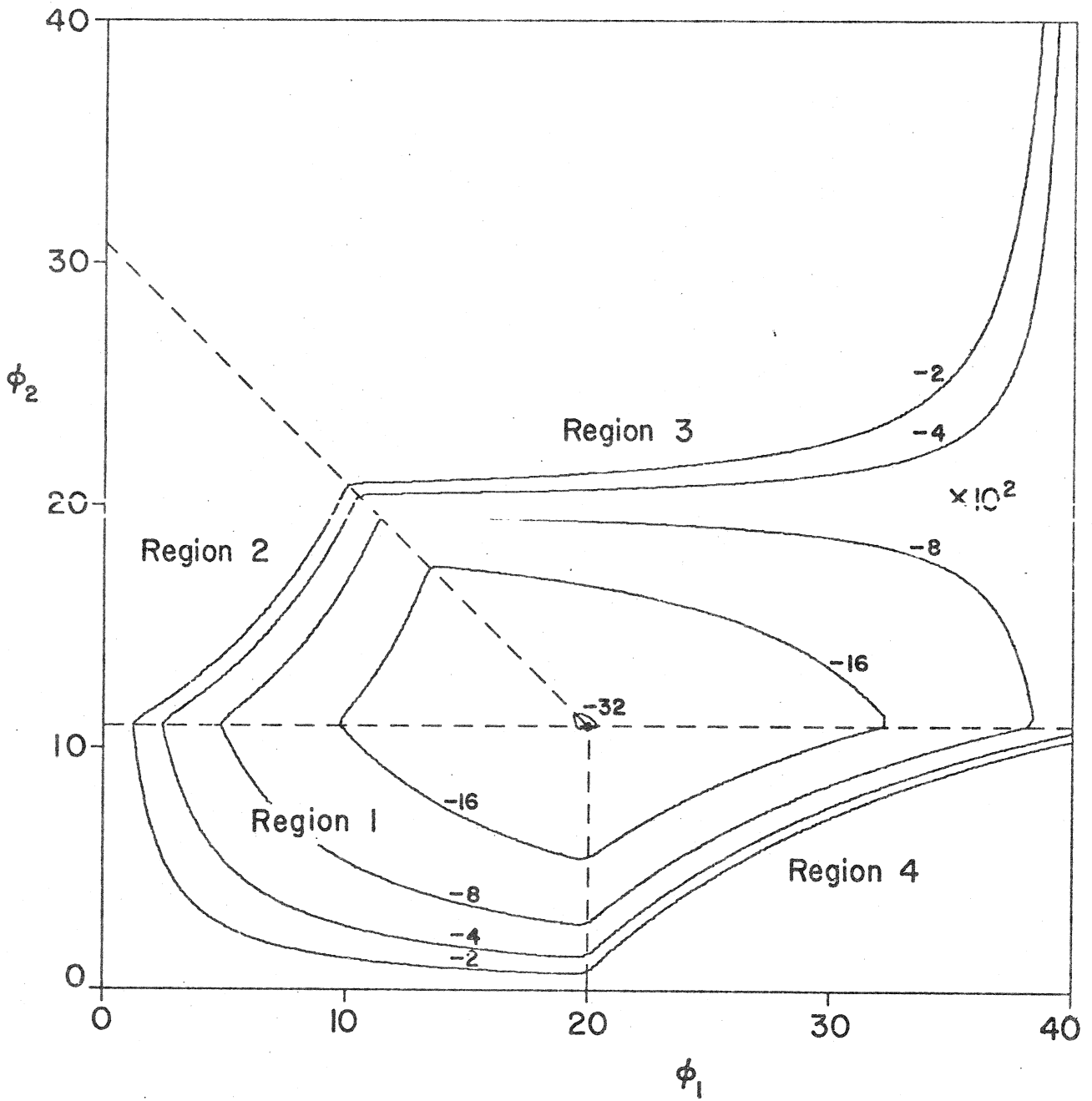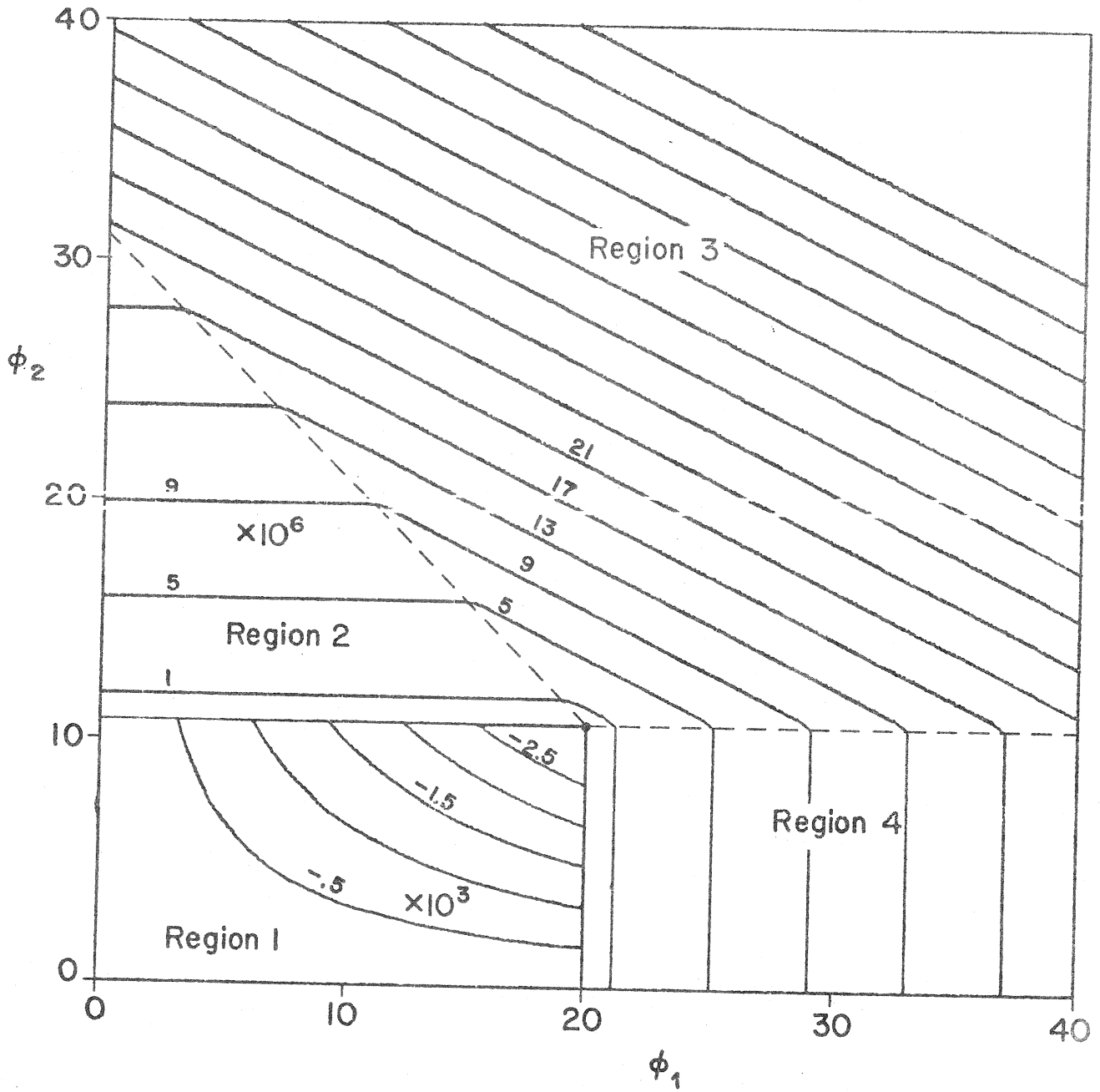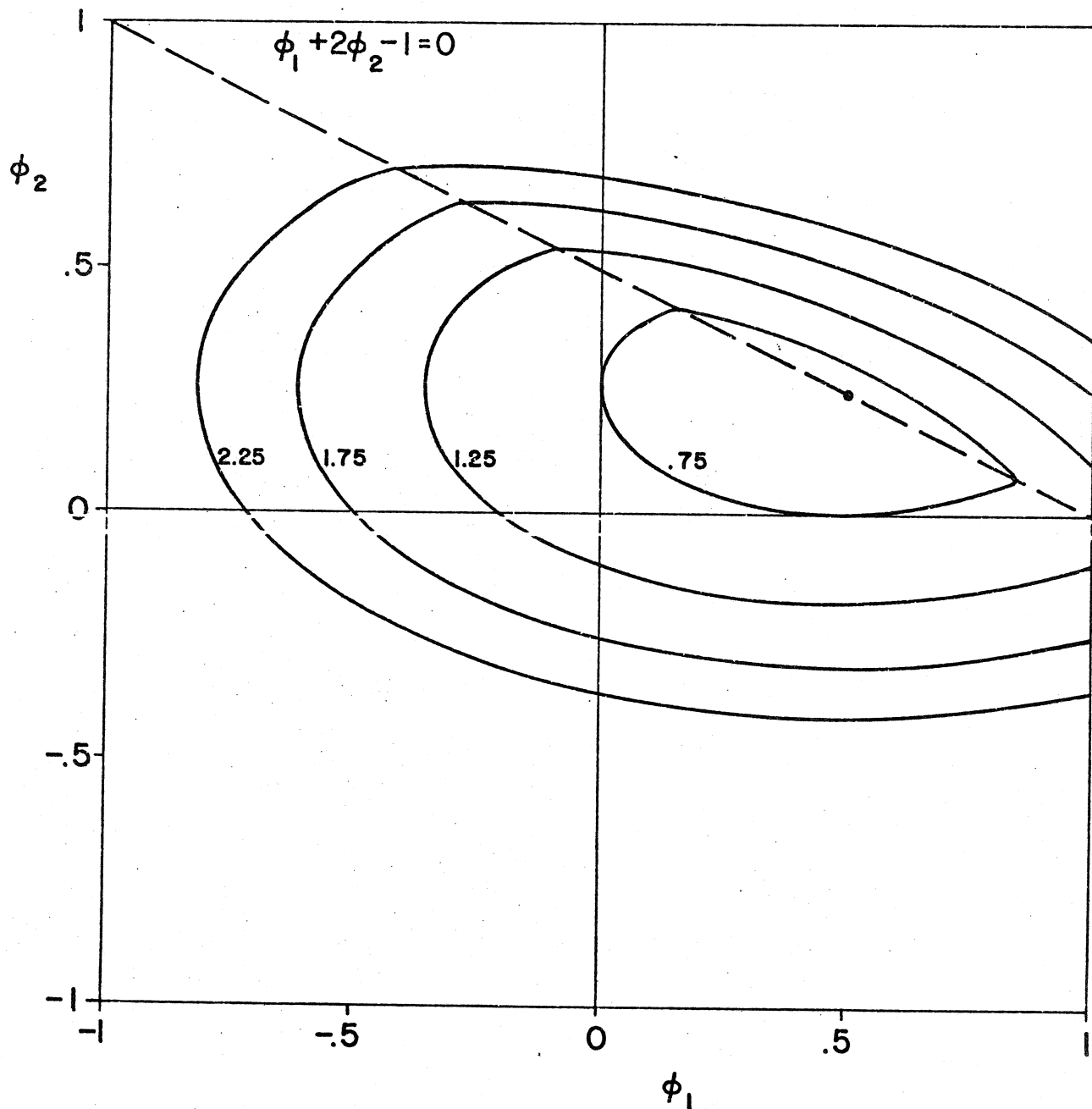Fig. 5. Contours for Quadratic Function when $\alpha=1$.

Fig. 1.

Fig. 2.

Fig. 3.

Fig. 4.

Fig. 5.