Comp. & Maths. with Appls. Vol. 11, No. 12, pp. 1165–1169, 1985
Printed in Great Britain.

0097-4943/85 $3.00 + .00

SOLUTION OF MINIMAX PROBLEMS USING EQUIVALENT DIFFERENTIABLE FUNCTIONS

REUVEN CHEN†
School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv, Israel

(*Received June* 1985)

Communicated by I. N. Katz

Abstract—A method is proposed for the solution of minimax optimization problems in which the individual functions involved are convex. The method consists of solving a problem with an objective function which is the sum of high powers or strong exponentials of the separate components of the original objective function. The resulting objective function, which is equivalent at the limit to the minimax one, is shown to be smooth as well as convex. Any efficient nonlinear programming method can be utilized for solving the equivalent problem. A number of examples are discussed.

1. INTRODUCTION

The minimax optimization problem is that of finding

$$\min_{\mathbf{x}} \{\max_{1 \le i \le n} [f_i(\mathbf{x})]\}, \tag{1}$$

where $\mathbf{x} = (x_1, \ldots, x_k)$. The functions $f_i(x)$ are assumed to be smooth, however, the main difficulty in solving (1) is usually related to the kinks in the objective function

$$F(\mathbf{x}) = \max_{1 \le i \le n} [f_i(\mathbf{x})].$$

These kinks are points at which $F(\mathbf{x})$ is not differentiable and in most cases, the solution point occurs at such a kink. The theory of nonlinear minimax has been thoroughly studied by Dem'yanov and Malozemov[9] who investigated the differentiability of the maximum function, and discussed the necessary and sufficient conditions for local and global solutions as well as the properties of the maximum problem. More recent theoretical work has been given by Ben-Tal and Zowe[3] and by Drezner[10]. The importance of the nonlinear minimax problem seems to exceed substantially the scope of solving problems which are initially of this nature since as shown by Bandler and Charalambous[2], any nonlinear programming problems with nonlinear constraints can be transformed into an equivalent unconstrained minimax problem. The numerical methods utilized for solving nonlinear minimax problems consisted mainly of approximating the $F(\mathbf{x}) = \max_{1 \le i \le n} [f_i(\mathbf{x})]$ function by close enough functions in which the kinks are smoothed out. Charalambous and Bandler[4] suggested the use of

$$\left\{\sum_{i=1}^{n} [f_i(\mathbf{x})]^P\right\}^{1/P}$$

as long as $f_i(\mathbf{x})$ are all positive and P is a large enough positive number. The result is a smooth approximation for $\max_{1 \le i \le n} [f_i(\mathbf{x})]$. Intuitively, if $f_i(\mathbf{x})$ are all positive valued, raising each of them to a high power will emphasize the largest one as compared to the others so that if $f_N(\mathbf{x}) > f_i(\mathbf{x})\ \forall\ i \ne N$, then

$$\sum_{i=1}^{n} [f_i(\mathbf{x})]^P \approx [f_N(\mathbf{x})]^P.$$

†Work done while in residence in the Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455 U.S.A.

Taking the p-th root yields

$$\lim_{P\to\infty}\left\{\sum_{i=1}^{n}[f_i(\mathbf{x})]^P\right\}^{1/P} = \max_i\,[f_i(\mathbf{x})]. \tag{2}$$

For a more rigorous proof see [11].

In the case that some or all of the functions may attain negative values, one can easily find a constant number A such that $f_i(\mathbf{x}) + A$ are positive for all the functions. Furthermore, the values of $[f_i(\mathbf{x}) + A]^P$ may be very large numbers, unmanageable by the computer in hand. A number M can be chosen such that $\{[f_i(\mathbf{x}) + A]/M\}^P$ will not exceed the capability of the computer to handle large numbers.

Zang[12] has given an alternative of smoothing the objective functions by utilizing polynomials on both sides of the kinks.

In the present work we suggest an improvement over these methods by utilizing equivalent differentiable functions, the solutions of which should coincide at the limit with those of the original minimax problem. One of these methods which has been successfully utilized for minimax location problems[7] is basically a variation of that by Charalambous and Bandler[4] mentioned above. Another method is presented here for the first time and the results of some test runs are given. The main features of these methods are the simplicity of their use as well as the rather fast convergence of the iterative procedure.

2. THE METHOD OF "SUM OF POWERS"

Let us assume in this section that $f_i(\mathbf{x})$ are all positive valued; if some of them are negative, a value A is chosen as explained above, and the $f_i(\mathbf{x})$ discussed here are the previous $f_i(\mathbf{x}) + A$. By the use of (2) above, a good approximation for $\min_\mathbf{x} \max_i [f_i(\mathbf{x})]$ can be achieved by $\min_\mathbf{x} \{\Sigma_{i=1}^n [f_i(\mathbf{x})]^P\}^{1/P}$ for a large enough value of P. As pointed out by Zang[12], if all the $f_i(\mathbf{x})$ functions are convex, the minimax problem is also convex, and as shown below, so are the approximations considered. As large as P may be, raising this approximate function to the P-th power should leave the minimum point unchanged. We can therefore solve

$$\min_{\mathbf{x}} \sum_{i=1}^{n} [f_i(\mathbf{x})]^P \tag{3}$$

for large values of P. This has been done very successfully with "radial" functions which occur in location problems[6]. In these cases, the functions are $f_i(r)$ where r is the Euclidean distance between a demand point and a service facility. The functions $f_i(r)$ represent the (positive) costs involved. In the location problems, convergence was quite good even for large values of P. A better strategy, similar to that often taken in the penalty function approach in the solution of constrained nonlinear program problems, was the following. First solve the approximate problem (3) for a moderate value of P and then use the solution as a starting point for a better solution with a larger value of P. This yielded usually much faster convergence.

The values of $[f_i(\mathbf{x})]^P$ may be too large, which, as pointed out above, calls for a normalization, namely, division of all the $f_i(\mathbf{x})$ functions by a constant. If increasing values of P are used successively, the value of the constant may be changed so that no overflow is encountered.

3. THE METHOD OF "SUM OF EXPONENTIALS"

The approximation for $\max_i [f_i(\mathbf{x})]$ has been shown to become better as the power P in $\Sigma_{i=1}^n [f_i(\mathbf{x})]^P$ becomes larger. From a broader point of view, if we took $\Sigma_{i=1}^n \phi[f_i(\mathbf{x})]$, the approximation became better for "stronger" functions $\phi(\cdot)$. By "strong" one should understand here a more convex function (i.e., having a larger positive second derivative). Among other things, the utilization of a convex function is of importance since, according to a quite well-known theorem (see, e.g., Avriel [1, p. 74]), if $f(\mathbf{x})$ is a convex function on $\mathbf{R}^n$ and ψ is a nondecreasing proper convex function on $\mathbf{R}$, then $\psi[f(\mathbf{x})]$ is convex on $\mathbf{R}^n$. In our case, this also means that if $f_i(\mathbf{x})$ are all convex, the approximation of solving $\min_{\mathbf{x}} \Sigma_{i=1}^{n} \phi[f_i(\mathbf{x})]$ preserves the important property of convexity. One can, therefore, choose any "strong" function $\phi(\cdot)$ and the one that comes to mind is the exponential function. One can therefore solve

$$\min_{\mathbf{x}} \sum_{i=1}^{n} \exp\left[\alpha \cdot f_i(\mathbf{x})\right], \tag{4}$$

for rather large positive values of α.

To show that the solution of (4) really approximates that of (1), note that in the sum, there are k (k being unity or a larger positive integer) largest terms. If α is large enough, we have

$$\sum_{i=1}^{n} \exp\left[\alpha \cdot f_i(\mathbf{x})\right] \cong k \cdot \exp\left[\alpha f_{\max}(\mathbf{x})\right]. \tag{5}$$

Taking the inverse of the exponential function we get at the limit of high α

$$\lim_{\alpha\to\infty} \frac{1}{\alpha} \ln \left\{ \sum_{i=1}^{m} \exp\left[\alpha f_i(\mathbf{x})\right] \right\} = \lim_{\alpha\to\infty} \left\{ \frac{1}{\alpha} k \cdot \exp\left[\alpha f_{\max}(\mathbf{x})\right] \right\} = \lim_{\alpha\to\infty} \frac{1}{\alpha} \ln k + f_{\max}(\mathbf{x}). \tag{6}$$

As α goes to infinity, the first term tends to zero. Since $(1/\alpha) \ln(\cdot)$ is a monotonically increasing function, it attains its maximum where its argument is maximal, which completes the proof that the solution of (4) coincides with that of (1). In fact, the same steps can be followed for the proof with any other increasing convex function $\phi(\cdot)$ instead of the exponent.

An advantage in solving (4) rather than (3) is that it does not make any difference if $f_i(\mathbf{x})$ attain negative or positive values. Again, it is advisable to solve the problem first for a relatively small value of α and use the solution as a first guess when the value of α increases by, say, a factor of 10. Here too, for large values of α, $\exp[\alpha \cdot f_i(\mathbf{x})]$ may be too large and normalization is necessary. This can be done in each step, i.e., after choosing the value of α, evaluate the current values of $f_i(\mathbf{x})$ and choose a constant c such that no value of $\exp[\alpha \cdot f_i(\mathbf{x}) - c]$ exceeds the values that can be handled by the computer. The subtraction of c actually amounts to the division of each of the terms by e^c. Our approximation will thus be found by solving

$$\min_{\mathbf{x}} \sum_{i} \exp\left[\alpha \cdot f_i(x) - c\right] \tag{7}$$

for large enough value of α associated with a suitable choice of c.

4. NUMERICAL EXAMPLES

The two minimax problems first suggested by Charalambous and Bandler[4] and later solved also by Zang[12] have now been solved by the "sum of exponentials" method. In "problem 1" the following three functions in two variables are given

$$\left.\begin{aligned} f_1(\mathbf{x}) &= x_1^4 + x_2^2 \\ f_2(\mathbf{x}) &= (2 - x_1)^2 + (2 - x_2)^2 \\ f_3(\mathbf{x}) &= 2 \exp(x_2 - x_1) \end{aligned}\right\}. \tag{8}$$

The results are shown in Table 1. Since in this case the functions are only of two variables, the nonlinear programming method chosen was the Newton method, which turned out to be very easy to use as well as fast converging. In other problems, with more variables, other methods should be considered since the inversion of the Hessian could be very time-consuming in a problem with many variables. Charalambous and Bandler[4] used the Fletcher subroutine whereas Zang employed one of the quasi-Newton methods. The result, accurate to eight sig-

Table 1. Results for Problem 1

α	c	ϵ	number of iterations	x_1	x_2	f_1	f_2	f_3
10^{-1}	-200	1.0×10^{-4}	8	0.9853028	0.7061335	1.4410893	2.7037179	1.5128342
10^{0}	-198	3.16×10^{-5}	4	1.0239997	0.8727946	1.8613339	2.2231549	1.7193042
10^{1}	-180	1.0×10^{-5}	6	1.0090040	0.9761852	1.9894433	2.0302698	1.9354273
10^{2}	0.317	3.16×10^{-6}	7	1.0011237	0.9972978	1.9991053	2.0031655	1.9923629
10^{3}	1.80×10^{3}	1.0×10^{-6}	6	1.0001152	0.9997258	1.9999125	2.0003179	1.9992220
10^{4}	1.98×10^{4}	3.16×10^{-7}	6	1.0000115	0.9999725	1.9999913	2.0000318	1.9999220
10^{5}	2.0×10^{5}	1.0×10^{-7}	6	1.0000012	0.9999973	1.9999991	2.0000032	1.9999922
10^{6}	2.0×10^{6}	3.16×10^{-8}	5	1.0000001	0.9999997	1.9999999	2.0000003	1.9999992
10^{7}	2.0×10^{7}	1.0×10^{-8}	5	1.0000000	1.0000000	2.0000000	2.0000000	1.9999999
10^{8}	2.0×10^{8}	3.16×10^{-9}	4	1.0000000	1.0000000	2.0000000	2.0000000	2.0000000

nificant figures, has been reached in 57 steps which have taken 76 msec CPU time. Using the more sophisticated nonlinear programming methods probably would have reduced the number of iterations, but would have made each step more time-consuming and would also make the programming much more complicated.

In the same way as in the previous works, the iteration started at the initial point (2.0, 2.0). As mentioned above, the procedure started with a small value of α (= 0.1) which was increased every time the termination criterion was satisfied. In the first stages, it would not have been beneficial to go for a very strict termination criterion because anyway, we are far from the final solution. Thus, for $\alpha = 0.1$ we started with $\epsilon = 10^{-4}$ where the termination criterion for each step was $|\Delta x_1| + |\Delta x_2| < \epsilon$. The values of α were increased by factors of 10 up to 10^8 whereas the values of ϵ were decreased by factors of $\sqrt{10}$ down to 3.16×10^{-9}. The computation was performed on the University of Minnesota Cyber 845 CDC computer.

In problem 2, solved in [4] and [12], the second and third functions are the same whereas the first is given by

$$f_1(\mathbf{x}) = x_1^2 + x_2^4. \tag{9}$$

The results are shown in Table 2. To attain the same accuracy of eight significant figures, starting from the same initial point (2.0, 2.0), a total of 43 steps were needed and the CPU time used was 59 msec.

Table 2. Results for Problem 2

α	c	ϵ	number of iterations	x_1	x_2	f_1	f_2	f_3
10^{-1}	-200	1.0×10^{-4}	8	1.2763344	0.7269008	1.9082413	2.1444299	1.1545785
10^{0}	-198	3.16×10^{-5}	4	1.2284402	0.7975116	1.9135908	2.0412836	1.2998128
10^{1}	-180	1.0×10^{-5}	6	1.1403082	0.8933298	1.9371676	1.9637889	1.5623142
10^{2}	-4.66	3.16×10^{-6}	5	1.1386056	0.8993522	1.9506358	1.9534259	1.5744336
10^{3}	1.75×10^{3}	1.0×10^{-6}	5	1.1389943	0.8995391	1.9520651	1.9523449	1.5741132
10^{4}	1.93×10^{4}	3.16×10^{-7}	4	1.1390333	0.8995579	1.9522086	1.9522365	1.5740809
10^{5}	1.95×10^{5}	1.0×10^{-7}	4	1.1390372	0.8995597	1.9522229	1.9522257	1.5740780
10^{6}	1.95×10^{6}	3.16×10^{-8}	4	1.1390376	0.8995599	1.9522243	1.9522246	1.5740777
10^{7}	1.95×10^{7}	1.0×10^{-8}	3	1.1390376	0.8995599	1.9522245	1.9522245	1.5740776

5. DISCUSSION

A new group of methods is suggested for the solution of the minimax problem (1). This consists of solving the problem

$$\min_{\mathbf{x}} \sum_{i=1}^{n} \phi[f_i(\mathbf{x})], \tag{10}$$

where ϕ is an increasing convex "strong" function. Problem (10) is a nonlinear optimization problem which may be solved in many ways. In the present work, the examples given are of functions with two variables only, therefore the simple Newton method could be applied. Another nonlinear programming method, tailored for the radial functions occurring in location problems has been successfully used by Chen[6,7]. This included the steepest descent method with an optimal step size easily found for this class of functions.

The present method has been found to be extremely simple for computation since the smoothing out of the kinks has been achieved by utilizing the exponential function rather than by a polynomial smoothing at each step.

Similar considerations can be utilized for other classes of problems. For example, the maximin problem

$$\max_{\mathbf{x}} \min_{i} [f_i(\mathbf{x})] \tag{11}$$

is interesting in location problems, where $f_i(\mathbf{x}) = w_i r_i$ and where $w_i r_i$ is the weighted Euclidean distance between two points. This is used for the location of obnoxious facilities such as nuclear reactors. Using similar arguments to those before, an equivalent problem is

$$\min_{\mathbf{x}} \sum_{i=1}^{n} \exp\left[-\alpha f_i(\mathbf{x}) + c\right] \tag{12}$$

with large enough value of α. Unfortunately, both (11) and (12) are not convex problems, and therefore, using any nonlinear programming method starting from any initial point would yield a local solution which is, certainly, not necessarily global. Finally, considerations similar to the "sum of powers" have been used for solving another class of problems, namely, minisum and minimax location-allocation problems[8]. Here too, the original problems as well as the approximations, are not necessarily convex and therefore the minima reached not global.

REFERENCES

1. M. Avriel, *Nonlinear Programming Analysis and Methods*. Prentice-Hall, Englewood Cliffs, NJ (1976).
2. J. W. Bandler and C. Charalambous, Nonlinear programming using minimax techniques. *J. Optim. Theory and Appl.* **13**, 607–619 (1976).
3. A. Ben-Tal and J. Zowe, Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems. *Math. Prog.* **24**, 70–91 (1982).
4. C. Charalambous and J. W. Bandler, Nonlinear minimax optimization as a sequence of least P-th optimization with finite values of P, *Int. J. Systems Sci.* **7**, 377–391 (1976).
5. J. Chatelton, D. Hearn and T. J. Love, A subgradient algorithm for certain minimax and minisum problems—the constrained case. *SIAM J. Control and Optim.* **20**, 455–469 (1982).
6. R. Chen, Solution of location problems with radial cost functions, *Comp. Math. with Appl.* **10**, 87–94 (1984).
7. R. Chen, Location problems with costs being sums of powers of Euclidean distances. *Comput. and Oper. Res.* **11**, 285–294 (1984).
8. R. Chen, Solution of minisum and minimax location-allocation problems with Euclidean distances. *Naval Res. Log. Quart.* **30**, 449–459 (1983).
9. V. Dem'yanov and V. N. Malozemov, On the theory of nonlinear minimax problems. *Russian Math. Surveys* **26**, 57–115 (1971).
10. Z. Drezner, On minimax optimization problems. *Math. Prog.* **22**, 227–230 (1982).
11. G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*. Cambridge University Press (1934).
12. I. Zang, A smoothing-out technique for min-max optimization. *Math. Prog.* **19**, 61–77 (1980).