INTERNAL REPORTS IN

SIMULATION, OPTIMIZATION AND CONTROL

No. SOC-12

NEW DIRECTIONS IN NONLINEAR PROGRAMMING FOR CIRCUIT DESIGN

J.W. Bandler

June 1973

FACULTY OF ENGINEERING
McMASTER UNIVERSITY
HAMILTON, ONTARIO, CANADA

Sixteenth Midwest Symposium on Circuit Theory

University of Waterloo
Waterloo, Ontario, Canada
April 12 - 13, 1973

Proceedings, vol. 1

NEW DIRECTIONS IN NONLINEAR PROGRAMMING FOR CIRCUIT DESIGN

John W. Bandler
Department of Electrical Engineering
McMaster University
Hamilton, Ontario, Canada.

Invited Paper

Abstract

This paper surveys some recent and important results in nonlinear programming useful in the efficient optimal design of circuits. Huang's general algorithm for unconstrained minimization is reviewed. It is also shown how constrained minimax problems can be solved exactly as unconstrained minimax problems, and then approximately solved using unconstrained gradient methods.

INTRODUCTION

In the area of numerical optimization, the past decade has seen a proliferation of *ad hoc* algorithms for unconstrained minimization of nonlinear functions, intuitive techniques for dealing with nonlinear constraints and "good-enough-for-practical-purposes" approaches to solving approximation problems associated with the frequency-or time-domain responses of circuits and systems. The trouble with all this is that it gives the impression that optimization is an art not a science. Furthermore, it leads to inefficient use of available computing resources, and impedes the advance of the area since it provides the climate of opinion that currently available techniques are already good enough.

The view which is held in certain quarters among engineers that more effort should be devoted to "real" problems rather that to computing techniques has also not helped the area of numerical methods. Presumably,

---

This work was supported under grant A7239 from the National Research Council of Canada.

their solution to solving some of these "real" problems when their programs consume large quantities of time and memory is to purchase a bigger and faster computer.

As will be gathered from the opening paragraphs one of the purposes of this paper is to discourage inefficiency in computing methods. To this end, a few recent and important results in the field of nonlinear programming will be highlighted. All the topics covered will be found useful in the optimal computer-aided design of circuits and also systems.

Huang's general algorithm for unconstrained minimization [1] is reviewed. The updating formulas employed by many of the efficient gradient optimization methods, notably the Fletcher-Powell [2] and Fletcher [3] methods fall out as special cases of Huang's algorithm.

This paper also indicates a number of ways in which constrained minimax approximation problems can be solved. Indeed, it will be apparent how any suitable algorithm for general nonlinear unconstrained minimax approximation, nonlinear programming, least pth approximation or unconstrained optimization can be used to solve both minimax approximation and nonlinear programming problems sufficiently well for the most exacting engineering purposes.

HUANG'S GENERALIZED ALGORITHM [1]

Suppose we have the problem of minimizing the unconstrained differentiable objective function U of a k-element column vector $\underset{\sim}{\phi}$. Let us denote the gradient vector of U by $\underset{\sim}{\nabla} U$, where

$$\underset{\sim}{\nabla} \triangleq \begin{bmatrix} \frac{\partial}{\partial \phi_1} \\ \frac{\partial}{\partial \phi_2} \\ \vdots \\ \frac{\partial}{\partial \phi_k} \end{bmatrix} \tag{1}$$

and the matrix of second partial derivatives by $\underset{\sim}{G}$ where

$$\underset{\sim}{G} \triangleq \underset{\sim}{\nabla}\, \underset{\sim}{\nabla}^T U \tag{2}$$

Superscript j will denote the jth iteration.

The updating formula derived by Huang for the approximation to $\underset{\sim}{G}^{-1}$ is

$$\underset{\sim}{H}^{j+1} = \underset{\sim}{H}^{j} + \rho \frac{\underset{\sim}{\delta}^{j}(C_1^j \underset{\sim}{\delta}^j + C_2^j \underset{\sim}{H}^{j^T} \underset{\sim}{\gamma}^j)^T}{(C_1^j \underset{\sim}{\delta}^j + C_2^j \underset{\sim}{H}^{j^T} \underset{\sim}{\gamma}^j)^T \underset{\sim}{\gamma}^j} - \frac{\underset{\sim}{H}^{j} \underset{\sim}{\gamma}^{j}(K_1^j \underset{\sim}{\delta}^j + K_2^j \underset{\sim}{H}^{j^T} \underset{\sim}{\gamma}^j)^T}{(K_1^j \underset{\sim}{\delta}^j + K_2^j \underset{\sim}{H}^{j^T} \underset{\sim}{\gamma}^j)^T \underset{\sim}{\gamma}^j} \tag{3}$$

where $\underset{\sim}{H}^j$ need not be symmetric; ρ, C_1^j, C_2^j, K_1^j and K_2^j are scalars,

$$\underset{\sim}{\delta}^j = \underset{\sim}{\phi}^{j+1} - \underset{\sim}{\phi}^j \tag{4}$$

taken in the direction $-\underset{\sim}{H}^{j^T} \underset{\sim}{\nabla} U^j$, and

$$\underset{\sim}{\gamma}^j = \underset{\sim}{\nabla} U^{j+1} - \underset{\sim}{\nabla} U^j \tag{5}$$

Huang's algorithm is based on conjugate search directions defined with respect to a quadratic model for U. Thus, it terminates in at most k iterations on a quadratic function, i.e.,

$$\underset{\sim}{\nabla U}^k = \underset{\sim}{0} \qquad (6)$$

Furthermore,

$$\underset{\sim}{H}^k = \rho \underset{\sim}{G}^{-1} \qquad (7)$$

Assume $\underset{\sim}{G}$ is positive definite. Then if ρ is positive, $\underset{\sim}{H}^k$ is positive definite. If ρ is zero, $\underset{\sim}{H}^k$ is the null matrix. If ρ is negative, $\underset{\sim}{H}^k$ is negative definite. Depending on the values of ρ, C_1^j, C_2^j, K_1^j, and K_2^j we can derive particular algorithms, including some well-known ones, as follows.

Let $\rho=1$, $C_1^j=1$, $C_2^j=0$, $K_1^j=0$ and $K_2^j=1$. Then

$$\underset{\sim}{H}^{j+1} = \underset{\sim}{H}^j + \frac{\underset{\sim}{\delta}^j \underset{\sim}{\delta}^{j^T}}{\underset{\sim}{\delta}^{j^T} \underset{\sim}{\gamma}^j} - \frac{\underset{\sim}{H}^j \underset{\sim}{\gamma}^j \underset{\sim}{\gamma}^{j^T} \underset{\sim}{H}^j}{\underset{\sim}{\gamma}^{j^T} \underset{\sim}{H}^j \underset{\sim}{\gamma}^j} \qquad (8)$$

which is the formula derived by Fletcher and Powell [2]. If $\underset{\sim}{H}^0$ is taken as a positive definite symmetric matrix then $\underset{\sim}{H}^j$ is also a positive definite symmetric matrix.

Let $\rho=1$,

$$\frac{C_2^j}{C_1^j} = \frac{-\underset{\sim}{\delta}^{j^T} \underset{\sim}{\gamma}^j}{\underset{\sim}{\delta}^{j^T} \underset{\sim}{\gamma}^j + \underset{\sim}{\gamma}^{j^T} \underset{\sim}{H}^j \underset{\sim}{\gamma}^j}$$

$K_1^j=1$ and $K_2^j=0$. Then

$$\underset{\sim}{H}^{j+1} = \underset{\sim}{H}^j - \frac{\underset{\sim}{\delta}^j \underset{\sim}{\gamma}^{j^T} \underset{\sim}{H}^j}{\underset{\sim}{\delta}^{j^T} \underset{\sim}{\gamma}^j} - \frac{\underset{\sim}{H}^j \underset{\sim}{\gamma}^j \underset{\sim}{\delta}^{j^T}}{\underset{\sim}{\delta}^{j^T} \underset{\sim}{\gamma}^j} + \left(1 + \frac{\underset{\sim}{\gamma}^{j^T} \underset{\sim}{H}^j \underset{\sim}{\gamma}^j}{\underset{\sim}{\delta}^{j^T} \underset{\sim}{\gamma}^j}\right) \frac{\underset{\sim}{\delta}^j \underset{\sim}{\delta}^{j^T}}{\underset{\sim}{\delta}^{j^T} \underset{\sim}{\gamma}^j} \qquad (9)$$

which is a formula derived by Fletcher [3], Broyden [4] and Goldfarb [5]. It has the same properties as the Fletcher-Powell formula.

Following the work of Dixon [6] it can be shown that the directions of search depend only on the values of $\underset{\sim}{H}^0$ and ρ. They are independent of the parameters C_1^j, C_2^j, K_1^j and K_2^j. As a consequence, if sequences of points $\underset{\sim}{\phi}^j$ are generated by formulas belonging to Huang's family for the same general nonquadratic function then the necessary and sufficient conditions for these sequences to be the same is that the formulas have the same values of $\underset{\sim}{H}^0$ and ρ.

Many algorithms in use belong to Huang's family with $\rho=1$ and they usually begin with $\underset{\sim}{H}^0$ equal to the unit matrix. If full linear searches are used to locate the minima in the search directions then, in theory, the same sequences of points should be generated by these algorithms. In practice, computational differences will mean that the sequences will not be exactly the same.

In view of these facts it is not surprising that there was little improvement in the area of unconstrained optimization between 1963 and 1970. Furthermore, since the algorithm proposed, for example, by Fletcher [3] abandons the full linear search (required for maintaining the quadratic termination property) and yet is remarkably efficient it is, in retrospect, also not surprisng why a number of users of other optimization methods have managed to obtain good results without spending a great deal of computing time doing cubic interpolation in the one-dimensional subproblems.

NONLINEAR APPROXIMATION, NONLINEAR PROGRAMMING AND UNCONSTRAINED OPTIMIZATION

Nonlinear Programming

Consider the problem of minimizing $U(\underset{\sim}{\phi})$ subject to $g_i(\underset{\sim}{\phi}) \geq 0$, $i=1,2,\ldots,m$, where m is the number of constraints. In general, $g_i(\underset{\sim}{\phi})$ will be taken as nonlinear and differentiable. It has been proved by Bandler and Charalambous [7,8] that a point satisfying the necessary conditions for optimality of the unconstrained function

$$V(\underset{\sim}{\phi}, \underset{\sim}{\alpha}) = \max_{1 \leq i \leq m} \left[U(\underset{\sim}{\phi}),\ U(\underset{\sim}{\phi}) - \alpha_i\, g_i(\underset{\sim}{\phi}) \right] \tag{10}$$

where

$$\underset{\sim}{\alpha} \triangleq [\alpha_1\ \alpha_2\ \ldots\ \alpha_m]^T \tag{11}$$

$$\alpha_i > 0 \quad i=1,2,\ldots,m \tag{12}$$

and sufficiently large will also satisfy the necessary conditions for optimality of the original nonlinear programming problem.

The implications of this are that any suitable method for non-linear minimax approximation can also be used to solve a nonlinear programming problem. The sufficiently large α_i will not, in general, be known in advance. Threshold values will, therefore, have to be estimated through a sequences of optimization problems.

Equality constraints can also be handled in this way. It is important to note, furthermore, that a feasible starting point does not have to be found. Indeed, there is no need to distinguish between feasible and nonfeasible points until a particular solution is obtained.

Constrained Minimax Approximation I

Consider the problem of minimizing $\max_{1 \le i \le n} f_i(\underset{\sim}{\phi})$ subject to $g_i(\underset{\sim}{\phi}) \ge 0$, $i=1,2,\ldots,m$. Here, n is the number of nonlinear, differentiable functions $f_i(\underset{\sim}{\phi})$. This problem can be recast as one of minimizing a new independent variable ϕ_{k+1} subject to

$$\phi_{k+1} - f_i(\underset{\sim}{\phi}) \ge 0 \quad i=1,2,\ldots,n \tag{13}$$

$$g_i(\underset{\sim}{\phi}) \ge 0 \quad i=1,2,\ldots,m \tag{14}$$

which is in the form of a conventional nonlinear programming problem.

This problem can again be reformulated as one of minimizing

$$V(\underset{\sim}{\phi},\underset{\sim}{\alpha}) = \max_{\substack{1 \le i \le n \\ 1 \le j \le m}} [\phi_{k+1}, \phi_{k+1} - \alpha_i(\phi_{k+1} - f_i(\underset{\sim}{\phi})), \phi_{k+1} - \alpha_{n+j}\, g_i(\underset{\sim}{\phi})] \tag{15}$$

where

$$\underset{\sim}{\alpha} = [\alpha_1 \; \alpha_2 \; \ldots \; \alpha_{n+m}]^T \tag{16}$$

One then proceeds as in the previous section [9].

Constrained Minimax Approximation II

We are given the same original problem as in the previous subsection. This time we consider the minimization of

$$W(\underset{\sim}{\phi},\underset{\sim}{w}) = \max_{\substack{1 \le i \le n \\ 1 \le j \le m}} [f_i(\underset{\sim}{\phi}), - w_j\, g_j] \tag{17}$$

which is taken as unconstrained and where

$$\underset{\sim}{w} \triangleq [w_1 \; w_2 \; \ldots \; w_m]^T \tag{18}$$

$$w_j > 0 \quad j=1,2,\ldots,m \tag{19}$$

It is assumed that $\max_{1 \leq i \leq n} f_i(\underset{\sim}{\phi}) > 0$ implies that certain design specifications have been ~~exceeded~~ violated and that $\max_{1 \leq i \leq n} f_i(\underset{\sim}{\phi}) < 0$ implies that they are satisfied. In this case, comparison with violated and satisfied constraints would be appropriate [9,10], and the resulting solution would be meaningful in the engineering sense since tradeoffs between response specifications and design constraints would be obtained.

Least pth Approximation

A suitable way of obtaining a good approximation to the unconstrained minimax problems formulated in this paper is to use one of the approaches suggested by Bandler and Charalambous [11,12]. Basically one can either use least pth approximation with very large values of p [11] or one can carry out a suitable sequence of least pth optimization problems with low values of p [12]. In both cases, of course, Fletcher's algorithm, for example, can be used effectively.

CONCLUSIONS

References are appended which will lead the interested person to papers which apply or extend the ideas presented here [13-16]. Attention should be drawn, for example, to the work of Charalambous who developed a family of minimization algorithms based on homogeneous models [16], out of which the Jacobson-Oksman algorithm [17] falls as a special case.

ACKNOWLEDGEMENT

The author is indebted to C. Charalambous, without whose assistance much of the work described and referred to in this paper would not have been possible.

REFERENCES

[1] H.Y. Huang, "Unified approach to quadratically convergent algorithms for function minimization", J. Optimization Theory and Applications, vol. 7, pp. 405-423, 1970.

[2] R. Fletcher and M.J.D. Powell, "A rapidly convergent descent method for minimization", Computer J., vol. 6, pp. 163-168, 1963.

[3] R. Fletcher, "A new approach to variable metric algorithms", Computer J., vol. 13, pp. 317-322, 1970.

[4] C.G. Broyden, "Quasi-Newton methods and their application to function minimization", Mathematics of Computation, vol. 21, pp. 368-381, 1967.

[5] D. Goldfarb, "A family of variable metric methods derived by variational methods", Mathematics of Computation, vol. 24, pp. 23-26, 1970.

[6] L.C.W. Dixon, "Variable metric algorithms: Necessary and sufficient conditions for identical behaviour of nonquadratic functions", J. Optimization Theory and Applications, vol. 10, pp. 34-40, 1972.

[7] J.W. Bandler and C. Charalambous, "A new approach to nonlinear programming", Proc. 5th Hawaii Int. Conf. on System Sciences, Honolulu, Hawaii, pp. 127-129, 1972.

[8] J.W. Bandler and C. Charalambous, "Nonlinear programming using minimax techniques", J. Optimization Theory and Applications, to be published.

[9] J.W. Bandler and T.V. Srinivasan, "Constrained minimax optimization by grazor search", Proc. 6th Hawaii Int. Conf. on System Sciences, Honolulu, Hawaii, pp. 125-128, 1973.

[10] J.R. Popović, J.W. Bandler and C. Charalambous, "Two computer programs for least pth approximation", 16th Midwest Symp. on Circuit Theory, Waterloo, Canada, April 1973.

[11] J.W. Bandler and C. Charalambous, "Practical least pth optimization of networks", IEEE Trans. Microwave Theory Tech., vol. MTT-20, pp. 834-840, 1972.

[12] C. Charalambous and J.W. Bandler, "New algorithms for minimax optimization", Nonlinear Programming Symp., Washington, D.C., March 1973.

[13] J.W. Bandler, V.K. Jha, C.M. Kudsia and J.R. Popović, "A microwave network optimization program", IEEE Int. Microwave Symp., Boulder, Colo., June 1973.

[14] C. Charalambous and J.W. Bandler, "New algorithms for network optimization", IEEE Int. Microwave Symp., Boulder, Colo., June 1973.

[15] J.W. Bandler and B.L. Bardakjian, "Least pth optimization of recursive digital filters", IEEE Int. Symp. on Circuit Theory, Toronto, Canada, April 1973.

[16] C. Charalambous, "New algorithms for unconstrained optimization," 16th Midwest Symp. on Circuit Theory, Waterloo, Canada, April 1973.

[17] D.H. Jacobson and W. Oksman, "An algorithm that minimizes homogeneous functions of N variables in N+2 iterations and rapidly minimizes general functions", J. Mathematical Analysis and Applications, vol. 38, pp. 535-552, 1972.