

# Sequential Minimax Search for a Zero of a Convex Function

By O. Gross and S. M. Johnson

**1. Introduction.** Many methods have been proposed and used for the computation of real roots of equations in a single real variable. Notable among these is Newton's Tangent Method. However, this procedure is quite costly in many instances wherein the function of which the zero is desired is inordinately complicated; indeed, it is usually the case—polynomials and other simple functions excepted—that derivatives are more complicated than the functions themselves. This last circumstance constitutes the principal motivation of later methods, such as the one given in [1], that are based solely upon evaluations of the functions themselves. To our knowledge, none of these later methods are sequential minimax in character except for the bisection procedure, and this only in a restricted sense.

Additional motivation for our method is provided by the observation that once a root is located on a sufficiently small interval by a positive and negative reading, the function will be inclined to be convex or concave throughout the interval, at least for typical analytic functions.

**2. Description of the Procedure.** In what follows, we shall describe a numerical procedure for solving the following problem:

“We know initially a positive and a negative value of a function at two given points. The function is continuous and convex and is otherwise unknown but computable. Given an integer  $n > 0$ , how do we locate its root within an interval of minimum length in  $n$  steps, where a step consists in calculating the value of the function at any point we choose?” The question has no definite answer until we specify, for example, that our procedure be sequential minimax; i.e., at each step of the procedure we assume that the worst possible situation might occur from that point on in the light of our present information about the function, and proceed to evaluate the function at a point that hedges against all contingencies so as to *guarantee* no more than a fixed interval length at that stage.

We now describe the procedure cycle:

Suppose we are in the situation in which we know

$$f(a) = Y_a > 0 \quad \text{and} \quad f(b) = -Y_b, \quad Y_b > 0, \quad \text{where} \quad a < b;$$

further, the root is greater than  $S$ , where  $S$  is known and we have  $n$  more readings to make.

Then we have bracketed the root on the interval  $(S, W)$ , where

$$W = a + (b - a) \frac{Y_a}{Y_a + Y_b}.$$

If  $n = 0$ , the computation ceases and the values  $S, W$  are recorded.

If  $n > 0$ , calculate the value of  $f$  at  $x = S + (W - S) \rho_n(Y_b/Y_a)$ . (Graphs of  $\rho_n$  for  $n = 1, 2, 3, 4$  are included in Fig. 5 at the end of this paper.)

If  $f(x) = Y_x > 0$ , set  $a' = x, b' = b$  and

$$S' = x + (x - a)Y_x/(Y_a - Y_x).$$

If, however,  $f(x) = -Y_x < 0$ , set  $a' = a, b' = x$ , and if  $Y_x > Y_b$  set  $S' = S$ ; otherwise, set

$$S' = \max(S, x - (b - x)Y_x/(Y_b - Y_x)).$$

Finally, set  $n' = n - 1$ .

Then we are in the situation in which we know  $f(a') = Y_a > 0, f(b') = -Y_{b'}$ ,  $Y_{b'} > 0$ , where  $a' < b'$ ; we know that the root is greater than  $S'$ ; and we have  $n'$  more readings to make. This completes the cycle. (As the problem is stated, we initially have  $S = a$ .)

In the next section, we shall illustrate how this procedure works on a particular example.

REMARK 1. The foregoing procedure is an *approximation* to the actual minimax procedure. The theoretically correct procedure involves replacing the expression  $\rho_n(Y_b/Y_a)$  in the formula for  $x$  above by

$$\rho_n\left(\frac{S - a}{b - a}, \frac{Y_b}{Y_a}\right),$$

where  $\rho_n(S, Y)$  is defined in Sec. 4. Since the function  $\rho_n(S, Y)$  is relatively insensitive to  $S$  in our choice of  $\rho_n(S, Y)$  near the minimax, we feel that the approximation

$$\rho_n(0, Y) \approx \rho_n(S, Y)$$

is justified, and define  $\rho_n(Y) \equiv \rho_n(0, Y)$ . This approximation renders the procedure more adaptable to machine computation.

REMARK 2. Since the average digital computer has difficulty in reading graphs, in order to program the procedure one would first find suitable algebraic approximations to the graphs of  $\rho_n$  after the fashion of C. Hastings, Jr. [3].

**3. A Numerical Application—Comparison with the Bisection Technique.** Suppose we want to bracket a zero of a certain function  $f$  defined over the interval  $(0, 1)$ . We know that  $f$  is continuous and convex, and that  $f(0) = 1, f(1) = -1$ . As a simple example, let  $f(x)$  be given by  $f(x) = \max(-1, (x - 1/3)(x/2 - 3))$ . We intend to make three evaluations of the function. Referring to the graph of  $R_3(0, Y)$  in Fig. 6, with  $Y = Y_0/Y_1 = 1$ , we see that we can guarantee locating the root on an interval of length 0.01 times the length of the original interval. However, since the graphs represent the worst that can happen, we expect to do much better.

We proceed to calculate:

CYCLE No. 1. We have  $a = 0, b = 1, Y_a = 1, Y_b = 1, S = 0, n = 3$ , whence by our formula we obtain  $W = 0.5$ , and we have located the root on  $(0, 0.5)$ . Next,

we compute  $x = 0 + (0.5 - 0)\rho_3(1) = 0.5(0.148) = 0.074$ , and we find that  $f(0.074) = 0.76839 > 0$ , so that  $a' = 0.074$ ,  $b' = 1$ , and

$$s' = 0.074 + \frac{0.074(0.76839)}{1 - 0.76839} = 0.31950.$$

Finally,  $n' = 2$ .

CYCLE No. 2. Dropping primes on the new variables, we have  $a = 0.074$ ,  $b = 1$ ,  $Y_a = 0.76839$ ,  $Y_b = 1$ ,  $S = 0.31950$ ,  $n = 2$ , whence by our formula we obtain

$$W = 0.074 + \frac{0.926(0.76839)}{1.76839} = 0.47636,$$

and we have located the root on  $(0.31950, 0.47636)$ . Next, we compute

$$x = 0.31950 + 0.15686\rho_2\left(\frac{1}{0.76839}\right) = 0.31950 + (0.15686)(0.198) = 0.35066,$$

and we find that  $f(0.35066) = -0.04795 < 0$ , so that  $a' = 0.074$ ,  $b' = 0.35066$ ; and since  $Y_x < Y_b$ , we have

$$S' = \max\left(0.31950, 0.35066 - \frac{(1 - 0.35066)(0.04795)}{1 - 0.04795}\right) = 0.31950.$$

Finally,  $n' = 1$ .

CYCLE No. 3. Now we have  $a = 0.074$ ,  $b = 0.35066$ ,  $Y_a = 0.76839$ ,  $Y_b = 0.04795$ ,  $S = 0.31950$ ,  $n = 1$ , whence by our formula we get

$$W = 0.074 + \frac{(0.35066 - 0.074)(0.76839)}{0.76839 + 0.04795} = 0.33441,$$

and we have located the root on  $(0.31950, 0.33441)$ . Next, we compute

$$\begin{aligned} x &= 0.31950 + (0.33441 - 0.31950)\rho_1\left(\frac{0.04795}{0.76839}\right) \\ &= 0.31950 + (0.01491)\rho_1(0.624) = 0.32440, \end{aligned}$$

and we find that  $f(0.32440) = 0.02535 > 0$ , so that  $a' = 0.32440$ ,  $b' = 0.35066$ , and

$$S' = 0.32440 + \frac{(0.32440 - 0.074)(0.02535)}{0.76839 - 0.02535} = 0.33294.$$

Finally,  $n' = 0$ .

CYCLE No. 4. Finally, we have  $a = 0.32440$ ,  $b = 0.35066$ ,  $Y_a = 0.02535$ ,  $Y_b = 0.04795$ ,  $S = 0.33294$ ,  $n = 0$ , whence by our formula we get

$$W = 0.32440 + \frac{(0.35066 - 0.32440)(0.02535)}{0.02535 + 0.04795} = 0.33348,$$

and we have located the root on  $(0.33294, 0.33348)$ . Now  $n = 0$ , so the computation ceases and the interval  $(0.33294, 0.33348)$  is recorded.

REMARK. Using the bisection technique, in 3 readings we would find the bracketing intervals of lengths 1, 0.5, 0.25, 0.125. Taking convexity into account, by our

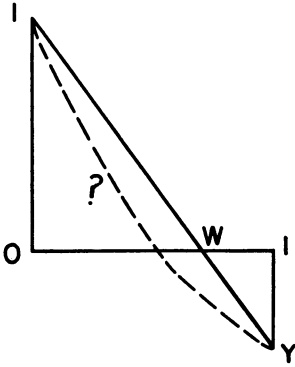


FIG. 1

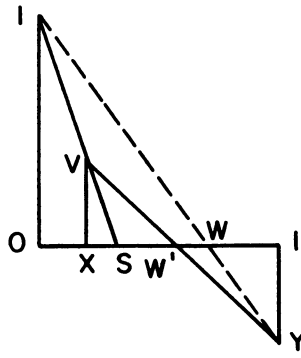


FIG. 2

method we obtain intervals of lengths 0.5, 0.15686, 0.01491, 0.00054, an obvious improvement.

**4. Derivation of the Functional Equation.** Given a continuous convex function  $f$  and two of its values, one positive and the other negative, (say  $f(x_1) > 0$  and  $f(x_2) < 0$ , with  $x_1 < x_2$ ), how should one search for the zero of the function that lies on the known interval  $(x_1, x_2)$ ? Using the principle of optimality of the theory of dynamic programming [2], we formulate the problem in terms of minimizing the maximum length of interval on which we can deduce that the zero is located in  $n$  readings taken sequentially.

First, by reduction to scale, we can always consider the diagram shown in Fig. 1, where  $f(0) = 1, f(1) = -Y, Y > 0, W = 1/(1 + Y)$ . Since  $f$  is convex, the zero must lie on  $(0, W)$ . If we have just one more reading to make ( $n = 1$ ), then we choose\*  $x$  on  $(0, W)$  and calculate  $f(x)$ . It can be shown by simple dominance arguments that no reading of  $f$  need ever be taken outside any interval on which the zero has been located. Having chosen  $x$  and calculated  $f(x)$ , we encounter exactly one of the following cases (barring  $f(x) = 0$ , of course, the best possible case).

CASE 1. If  $f(x) = V > 0$ , then by drawing lines joining known points on the graph of  $f$ , we have the picture shown in Fig. 2, with the root located on  $[S, W']$ , as implied by the convexity of  $f$ .

CASE 2. If  $f(x) = -V' < 0$ , the picture is shown in Fig. 3, with the root located on  $(\max(0, S'), W'')$  as indicated.

Let  $S'' = \max(0, S')$ . If  $x$  is the final point at which  $f(x)$  is to be determined (i.e.,  $n = 1$ ) then it is chosen so as to minimize the maximum possible value of  $\max(W' - S, W'' - S'')$  consistent with our choice.

Let us consider the general  $n$ -stage process in which we have several more readings to make. In either of the two cases diagrammed above, all the essential data can be described by a basic triangle determined by a two-parameter system as follows.

In Fig. 2, we can conclude that the graph of  $f$  lies above the line segment  $VS$  and below the segment  $VW'Y$  (where the letters stand for both points and values

\* The particular optimal choice will be derived in what follows.

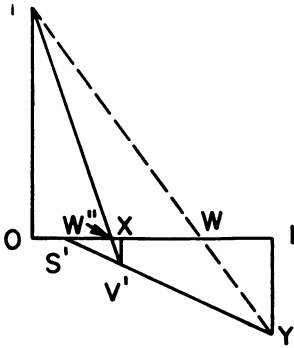


FIG. 3

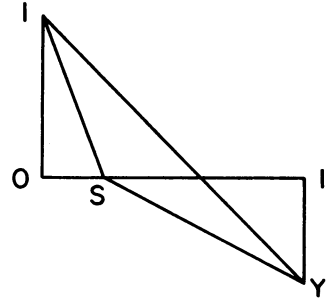


FIG. 4

in the obvious manner). If we now draw the line segment  $SY$ , it may or may not be crossed by the graph of  $f$ ; but if it is crossed at some point  $P$ , say, we can replace that portion of the curve lying below  $PY$  by  $PY$ . This device does not violate the convexity condition nor does it exclude the worst possible case. This can be shown by simple dominance arguments based on our knowledge of the function at any particular stage.

The essential data can thus be described by the triangle  $VS Y$ . Since a vertical reduction in scale leaves the problem invariant, and a horizontal reduction leaves it relatively invariant, the triangle  $VS Y$  can be replaced by a triangle of standard form described pictorially by two parameters  $Y, S$  (say), as shown in Fig. 4.

Similarly, in the second case (Fig. 3), the graph of  $f$  lies above  $S'V'$  and below  $1W''V'$ . Draw the line segment from 1 to  $S'$  and replace any portion of the graph of  $f$  lying below this line by the line itself from 1 to the point of crossing. This does not affect the choice of subsequent values of  $x$ , since they will all be chosen on minimal bracketing intervals. Again, this can be shown by simple dominance arguments. Thus, in the second case also by a suitable reduction to scale we are led to another representation of Fig. 4.

Now define

$R_n(S, Y)$  = the minimum length of interval on which we can guarantee locating the zero in  $[0, 1]$  of any convex function  $f$ , given that  $f(0) = 1, f(1) = -Y < 0$ , the root is  $> S$ , and we have  $n$  readings to make.

If  $n = 0$ , we clearly have

$$R_0(S, Y) = \frac{1}{1 + Y} - S.$$

Next, using the principle of optimality, and taking into account the scale factors, we obtain for  $n > 0$  the following recurrence relation:

$$R_n(S, Y) = \min_{S \leq x \leq 1/(1+Y)} \max \begin{cases} \max_{0 \leq V \leq 1-x/(1+Y)} (1-x)R_{n-1}\left(\frac{x}{1-x}, \frac{V}{1-V}, \frac{Y}{V}\right) \\ \max_{0 \leq V' \leq Y(x-S)/(1-S)} xR_{n-1}\left(\frac{xY-V'}{x(Y-V')}, V'\right) \end{cases}$$

with the upper and lower expressions after the brace corresponding to the first and

second cases, respectively. The scale factors are obtained in a completely elementary manner by means of similar triangles.

The ranges of the variables  $S, Y$  above are given by  $Y \geq 0, 0 \leq S \leq 1/(1 + Y)$ . To render the expressions more amenable to machine computation, we make the following substitutions:

$$W = \frac{1}{1 + Y}, \phi_n(S, W) = R_n(S, Y), \text{ whence } R_n(S, Y) = \phi_n\left(S, \frac{1}{1 + Y}\right).$$

An additional modification of the quantified variables  $V, V'$  finally reduces the system to  $\phi_0(S, W) = W - S$ ,

$$\phi_n(S, W) = \min_{s \leq x \leq w} \max \begin{cases} (1 - x) \max_{x \leq t' \leq w} \phi_{n-1}\left(\frac{t' - x}{1 - x}, \frac{W(t' - x)}{t' - Wx}\right) \\ x \max_{s \leq t \leq x} \phi_{n-1}\left(\frac{t}{x}, \frac{W(1 - t)}{W(1 - x) + x - t}\right) \end{cases}$$

where  $0 \leq S \leq W \leq 1$ .

The functions  $\phi_n(S, W)$  are then computed for  $n = 1, 2, 3, 4$  by means of a discrete approximation using various grid sizes and linear interpolation.

The minimizing values of  $x$  are recorded and these form the basis for our optimal policy; i.e.,  $x^* = x_n^*(S, W)$  is the point at which we evaluate our unknown function  $f$  given that the root lies on  $(S, W)$  in our basic triangle and there are  $n$  evaluations to be made. The point  $x_n^*$  is, of course, itself measured on the basic triangle. To take care of the general situation, we must relate our readings to the original scale. If we let  $\rho_n(S, Y)$  denote the fraction of the distance between  $S$  and  $W$  occupied by  $x_n^*$ , we readily obtain

$$\rho_n(S, Y) = \frac{x_n^*(S, W) - S}{W - S},$$

where  $W = 1/(1 + Y)$ . It is now a relatively easy matter to relate  $x_n^*$  to the original scale and thus to obtain the procedure cycle outlined in Sec. 2.

Graphs of the functions  $R_n(0, Y)$  and  $\rho_n(0, Y)$  for  $n = 1, 2, 3, 4$  are included at the end of this paper (Figs. 5 and 6).

**5. Remarks.** We shall close with a few remarks intended primarily to validate certain assumptions made, tacitly or otherwise, in the derivation of the functional equation.

1. Suppose we are at a certain stage in an optimal sequential minimax search. We have computed several values of  $f$  and are ready to choose our next point of evaluation. Let  $a, b$  denote the closest evaluation points on the left and right, respectively, of the minimal interval  $(S, W)$  on which the root is known to lie. Then in an optimal procedure, we choose  $x$  on  $(a, b)$ . To see this, we need only state that since the unknown function  $f$  may indeed be linear both to the left of  $a$  and to the right of  $b$  (a possible contingency), it is easy to see that any such subsequent reading would in such a contingency afford us no information regarding the character of  $f$  within  $(a, b)$  pertinent to the location of its zero, not already implied by the quantities  $a, b, f(a), f(b), S$  and  $W$ , or indeed by any of our previous readings.

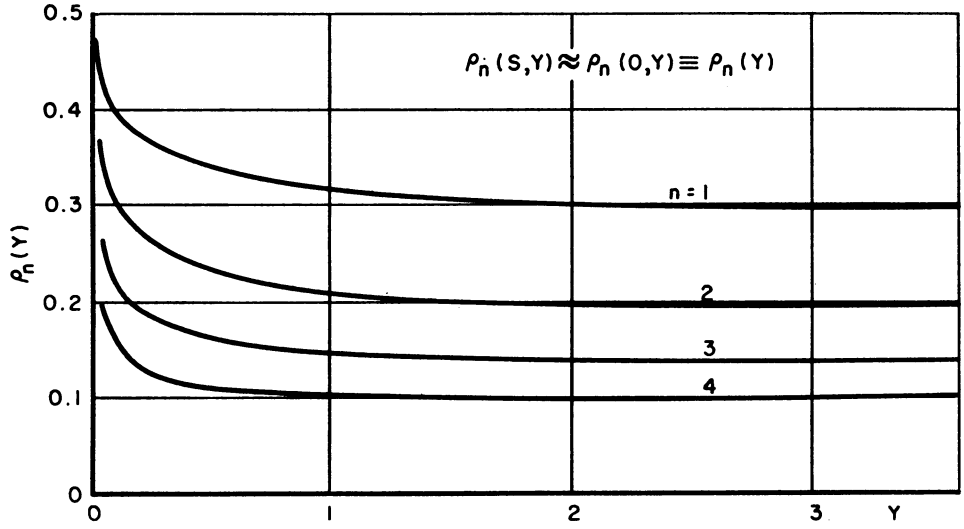


Fig. 5

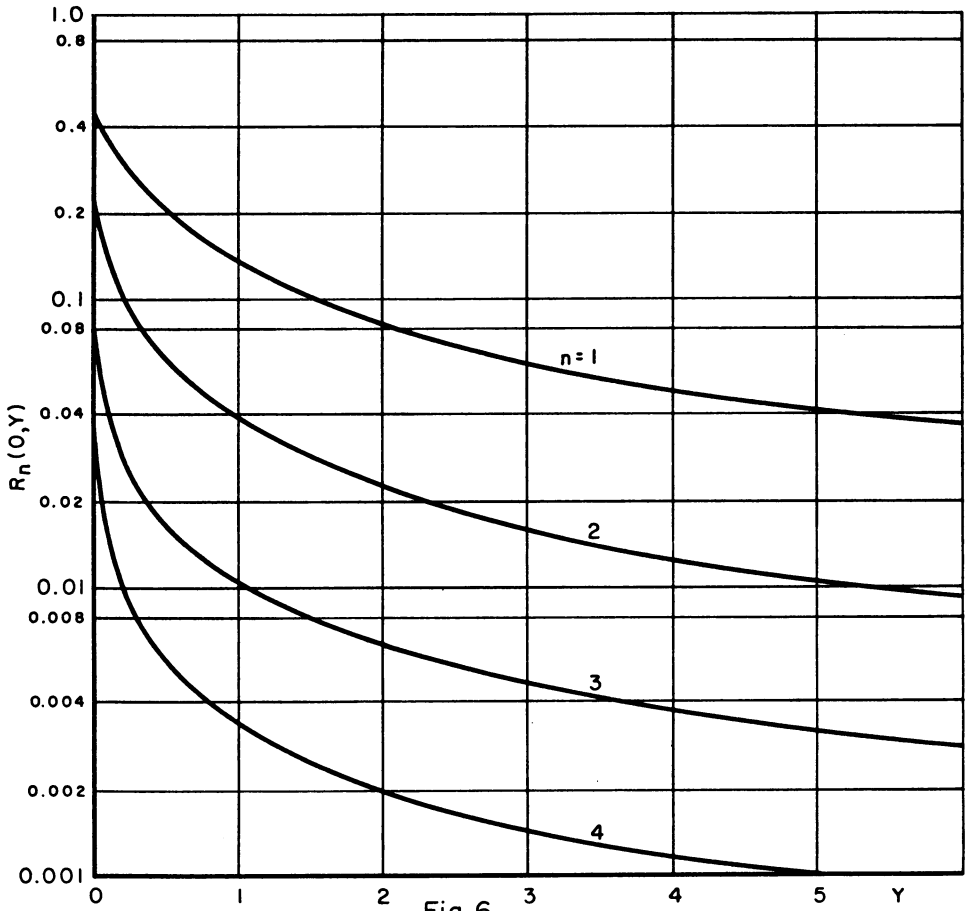


Fig. 6

Similar but slightly more involved arguments can be given to show that the next reading should be taken on the interval  $(S, W)$ .

2. In the treatment of case 2, it was tacitly assumed that  $Y_z < Y$ ; i.e., the  $S'Y$  line has a negative slope as shown in Fig. 3. Again, it can be shown by dominance arguments that the worst situation occurs when  $f$  is monotone, and indeed when the graph of  $f$  lies above the line  $SY$ . It is this condition that determines the limits of the quantified variable  $V$  in the upper line of the functional equation for  $R_n$ . It is precisely such dominance considerations as these that, though enabling us to express the functional equation in a relatively simple form and to obtain an optimal "policy" therefrom via its recursive computation, prevent us from obtaining an optimal "procedure" directly from the equation. However, all the contingencies, together with the information about the function  $f$  they afford, are taken care of in the procedure outlined in Sec. 2.

3. The function  $R_n(S, Y)$  is separately decreasing in  $S$  and  $Y$ . To see this for the first variable, for example, upon recalling the definition of  $R_n(S, Y)$ , we need only observe that the information that the root is  $> S$  includes the information that the root is  $> S'$  if  $S' < S$ . On the basis of the additional information, then, we can clearly *guarantee at least* as short a final bracketing interval with a larger  $S$  value if we are proceeding optimally; i.e., we have  $R_n(S, Y) \leq R_n(S', Y)$  if  $S > S'$ . An analogous argument applies for the second variable.

4. Finally, we note that an analytic treatment of the case  $n = 1$  provides a check on the machine computations.

Rand Corporation, Santa Monica, California

1. WILLIAM H. SHIELDS, "Determination of zeros of polynomials and real roots of transcendental equations," Report R-77, Instrumentation Lab., Mass. Inst. of Tech., May 1954.

2. RICHARD BELLMAN, "The theory of dynamic programming," Amer. Math. Soc., *Bull.*, v. 60, November, 1954, p. 503-515.

3. CECIL HASTINGS, JR., *Approximations for Digital Computers*, Princeton University Press, Princeton, New Jersey, 1955.