A CLASS OF OPTIMAL-ORDER ZERO-FINDING METHODS USING DERIVATIVE EVALUATIONS

Richard P. Brent

Computer Centre,
Australian National University,
Canberra, A.C.T. 2600, Australia

1. INTRODUCTION

It is often necessary to find an approximation to a simple zero ζ of a function f, using evaluations of f and f'. In this paper we consider some methods which are efficient if f' is easier to evaluate than f. Examples of such functions are given in Sections 5 and 6.

The methods considered are stationary, multipoint, iterative methods, "without memory" in the sense of Traub [64]. Thus, it is sufficient to describe how a new approximation (x_1) is obtained from an old approximation (x_0) to ζ. Since we are interested in the order of convergence of different methods, we assume that f is sufficiently smooth near ζ, and that x_0 is sufficiently close to ζ. Our main result is:

<u>Theorem 1.1</u>

There exist methods, of order 2ν, which use one evaluation of f and ν evaluations of f' for each iteration.

By a result of Meersman and Wozniakowski, the order 2ν is the highest possible for a wide class of methods using the same information (i.e., the same number of evaluations of f and f' per iteration): see Meersman [75]. The "obvious" interpolatory methods have order $\nu + 1$, but the optimal order 2ν may be obtained by evaluating f' at the correct points. These points are determined by some properties of orthogonal and "almost orthogonal" polynomials.

If $\nu + 1$ evaluations of f are used, instead of one function evaluation and ν derivative evaluations, then the optimal order is 2^ν for methods without memory (Kung and Traub [73,74], Wozniakowski [75a,b]), and $2^{\nu+1}$ for methods with memory (Brent, Winograd and Wolfe [73]). Thus, our methods are only likely to be useful for small ν or if f' is much cheaper than f.

Special Cases

Our methods for $\nu \geq 3$ appear to be new. The cases $\nu = 1$ (Newton's method) and $\nu = 2$ (a fourth-order method of Jarratt [69]) are well known. Our sixth-order method (with $\nu = 3$) improves on a fifth-order method of Jarratt [70].

Generalizations

Generalizations to methods using higher derivatives are possible. One result is:

Theorem 1.2

For $m > 0$, $n \geq 0$, and k satisfying $m + 1 \geq k > 0$, there exist methods which, for each iteration, use one evaluation of $f, f', \ldots, f^{(m)}$, followed by n evaluations of $f^{(k)}$, and have order of convergence $m + 2n + 1$.

The methods described here are special cases of the methods of Theorem 1.2 (take $k = m = 1$, and $\nu = n + 1$) . Since proof of Theorem 1.2 is given in Brent [75], we omit proofs here, and adopt an informal style of presentation. Other possible generalizations are mentioned in Section 7.

2. MOTIVATION

We first consider methods using one evaluation of f , and two of f' , per iteration. Let x_0 be a sufficiently good approximation to the simple zero ζ of f , $f_0 = f(x_0)$, and $f_0' = f'(x_0)$. Suppose we evaluate $f'(\tilde{x}_0)$, where

$$\tilde{x}_0 = x_0 - \alpha f_0/f_0' ,$$

and α is a nonzero parameter. Let $Q(x)$ be the quadratic polynomial such that

$$Q(x_0) = f_0 ,$$

$$Q'(x_0) = f_0' ,$$

and

$$Q'(\tilde{x}_0) = f'(\tilde{x}_0) ,$$

and let x_1 be the zero of $Q(x)$ closest to x_0 . Jarratt [69] essentially proved:

Theorem 2.1

$$x_1 - \zeta = O(|x_0 - \zeta|^\rho)$$

as $x_0 \to \zeta$, where

$$\rho = \begin{cases} 3 & \text{if } \alpha \neq 2/3 , \\ 4 & \text{if } \alpha = 2/3 . \end{cases}$$

Thus, we choose $\alpha = 2/3$ to obtain a fourth-order method. The proof of Theorem 2.1 uses the following lemma:

Lemma 2.1

If $P(x) = a + bx + cx^2 + dx^3$ satisfies

$$P(0) = P'(0) = P'(2/3) = 0 ,$$

then $P(1) = 0$.

Applying Lemma 2.1, we may show that (for $\alpha = 2/3$)

$$f(x_N) - Q(x_N) = O(\delta^4) ,$$

where

$$x_N = x_0 - f_0/f'_0$$

is the approximation given by Newton's method, and

$$\delta = |f_0/f'_0| = |x_N - x_0| .$$

Now

$$x_N - x_1 = O(\delta^2) ,$$

and

$$f'(x) - Q'(x) = O(\delta^2)$$

for x near x_N, so

$$|f(x_1)| = |f(x_1) - Q(x_1)|$$
$$\leqslant |f(x_N) - Q(x_N)| + |f'(\xi) - Q'(\xi)|\cdot|x_N - x_1|$$

for some ξ between x_N and x_1. Thus

$$|f(x_1)| = O(\delta^4) + O(\delta^2\cdot\delta^2) = O(\delta^4) ,$$

and

$$x_1 - \zeta = O(|f(x_1)|) = O(\delta^4) = O(|x_0 - \zeta|^4) .$$

3. A SIXTH-ORDER METHOD

To obtain a sixth-order method using one more derivative evaluation than the fourth-order method described above, we need distinct, nonzero parameters, α_1 and α_2, such that

$$P(0) = P'(0) = P'(\alpha_1) = P'(\alpha_2) = 0$$

implies $P(1) = 0$, for all fifth-degree polynomials

$$P(x) = a + bx + \dots + fx^5 .$$

Thus, we want the conditions

$$2\alpha_1 c + \dots + 5\alpha_1^4 f = 0$$

and

$$2\alpha_2 c + \dots + 5\alpha_2^4 f = 0$$

to imply

$$c + \dots + f = 0 .$$

Equivalently, we want

$$\operatorname{rank}\begin{bmatrix} 2\alpha_1 & 3\alpha_1^2 & 4\alpha_1^3 & 5\alpha_1^4 \\ 2\alpha_2 & 3\alpha_2^2 & 4\alpha_2^3 & 5\alpha_2^4 \\ 1 & 1 & 1 & 1 \end{bmatrix} = 2 ,$$

i.e.,

$$\operatorname{rank}\begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \alpha_1^3 \\ 1 & \alpha_2 & \alpha_2^2 & \alpha_2^3 \\ 1/2 & 1/3 & 1/4 & 1/5 \end{bmatrix} = 2 ,$$

i.e., for some w_1 and w_2,

(3.1) $$w_1\alpha_1^i + w_2\alpha_2^i = 1/(i + 2)$$

for $0 \leqslant i \leqslant 3$.

Since $1/(i + 2) = \int_0^1 x^i\cdot x\,dx$, we see from (3.1) that α_1 and α_2 should be chosen as the zeros of the Jacobi polynomial, $G_2(2, 2, x) = x^2 - 6x/5 + 3/10$, which is orthogonal to lower degree polynomials, with respect to the weight function x, on $[0, 1]$.

Let $y_i = x_0 - \alpha_i f_0/f'_0$, $x_N = x_0 - f_0/f'_0$, $\delta = |f_0/f'_0|$, and let $Q(x)$ be the cubic polynomial such that

$$Q(x_0) = f_0 , \quad Q'(x_0) = f'_0 ,$$

and

$$Q'(y_i) = f'(y_i)$$

for $i=1,2$. Then

$$f(x) - Q(x) = O(\delta^4)$$

for x between x_0 and x_N, but

$$f(x_N) - Q(x_N) = O(\delta^6) ,$$

because of our choice of α_1 and α_2 as zeros of $G_2(2, 2, x)$.

(This might be called "superconvergence": see de Boor and Swartz [73].)

A Problem

Since

$$x_N - x_1 = O(\delta^2)$$

and

$$f'(x) - Q'(x) = O(\delta^3)$$

for x near x_N, proceeding as above gives

$$|f(x_1)| = O(\delta^6) + O(\delta^3 \cdot \delta^2) = O(\delta^5) ,$$

so the method is only of order five, not six.

A Solution

After evaluating $f'(y_1)$, we can find an approximation $\tilde{x}_N = \zeta + O(\delta^3)$ which is (in general) a better approximation to ζ than is x_N. From the above discussion, we can get a sixth-order method if we can ensure superconvergence at $\tilde{x}_N$ rather than x_N. Define $\tilde{\alpha}_1$ by

$$\tilde{\alpha}_1(\tilde{x}_N - x_0) = \alpha_1(x_N - x_0) .$$

In evaluating f' at $y_1 = x_0 + \tilde{\alpha}_1(\tilde{x}_N - x_0)$, we effectively used $\tilde{\alpha}_1 = \alpha_1 + O(\delta)$ instead of α_1, so we must perturb α_2 to compensate for the perturbation in α_1.

From (3.1), we want $\tilde{\alpha}_2$ such that, for some $\tilde{w}_1$ and $\tilde{w}_2$,

$$\tilde{w}_1\tilde{\alpha}_1^i + \tilde{w}_2\tilde{\alpha}_2^i = 1/(i+2) \tag{3.2}$$

for $0 \le i \le 2$. Thus

$$\text{rank}\begin{bmatrix} 1 & \tilde{\alpha}_1 & \tilde{\alpha}_1^2 \\ 1 & \tilde{\alpha}_2 & \tilde{\alpha}_2^2 \\ 1/2 & 1/3 & 1/4 \end{bmatrix} = 2 ,$$

which gives

$$\tilde{\alpha}_2 = (3 - 4\tilde{\alpha}_1)/(4 - 6\tilde{\alpha}_1) = \alpha_2 + O(\delta) .$$

Since

$$\tilde{w}_j = w_j + O(\delta)$$

for $j=1,2$, we have

$$\tilde{w}_1\tilde{\alpha}_1^3 + \tilde{w}_2\tilde{\alpha}_2^3 = 1/5 + O(\delta) . \tag{3.3}$$

(Compare (3.1) with $i = 3$.) If we evaluate f' at $\tilde{y}_2 = x_0 + \tilde{\alpha}_2(\tilde{x}_N - x_0)$, and let x_1 be a sufficiently good approximation to the appropriate zero of the cubic which fits the data obtained from the f and f' evaluations, then (3.2) and (3.3) are sufficient to ensure that the method has order six after all.

4. METHODS OF ORDER 2ν

In this section we describe a class of methods satisfying Theorem 1.1. The special cases $\nu = 2$ and $\nu = 3$ have been given above.

It is convenient to define $n = \nu - 1$. The Jacobi polynomial $G_n(2, 2, x)$ is the monic polynomial, of degree n, which is orthogonal to all polynomials of degree $n - 1$, with respect to the weight function x, on $[0, 1]$. Let $\alpha_1,\ldots,\alpha_n$ denote the zeros of $G_n(2, 2, x)$ in any fixed order. We describe a class of methods of order $2(n + 1)$, using evaluations of $f(x_0)$, $f'(x_0)$, and $f'(y_1),\ldots,f'(y_n)$, where the points $y_1,\ldots,y_n$ are determined during the iteration.

The Methods

1. Evaluate $f_0 = f(x_0)$ and $f'_0 = f'(x_0)$.
2. If $f_0 = 0$ set $x_1 = x_0$ and stop, else set $\delta = |f_0/f'_0|$.
3. For $i=1,\ldots,n$ do steps 4 to 7.

4. Let p_i be the polynomial, of minimal degree, agreeing with the data obtained so far. Let z_i be an approximate zero of p_i, satisfying $z_i = x_0 + O(\delta)$ and $p_i(z_i) = O(\delta^{i+2})$. (Any suitable method, e.g. Newton's method, may be used to find z_i.)

5. Compute $\alpha_{i,j} = \alpha_{i-1,j}(z_{i-1} - x_0)/(z_i - x_0)$ for $j=1,\ldots,i-1$. (Skip if $i = 1$.)

6. Let q_i be the monic polynomial, of degree $n + 1 - i$, such that $\int_0^1 P(x)\, q_i(x) \left(\prod_{j=1}^{i-1} (x - \alpha_{i,j})\right) x\,dx = 0$ for all polynomials P of degree $n - i$. (The existence and uniqueness of q_i may be shown constructively: see Brent [75].) Let $\alpha_{i,i}$ be an approximate zero of q_i, satisfying $\alpha_{i,i} = \alpha_i + O(\delta)$ and $q_i(\alpha_{i,i}) = O(\delta^{i+1})$.

7. Evaluate $f'(y_i)$, where

$$y_i = x_0 + \alpha_{i,i}(z_i - x_0) .$$

8. Let p_{n+1} be as at step 4, and x_1 an approximate zero of p_{n+1}, satisfying $x_1 = x_0 + O(\delta)$ and $p_{n+1}(x_1) = O(\delta^{2n+3})$.

Asymptotic Error Constants

The asymptotic error constant of a stationary zero-finding method is defined to be

$$K = \lim_{x_0 \to \zeta} (x_1 - \zeta)/(x_0 - \zeta)^{\rho} ,$$

where ρ is the order of convergence. (Since ρ is an integer for all methods considered here, we allow K to be signed.) Let K_ν be the asymptotic error constant of the methods (of order 2ν) described above. The general form of K_ν is not known, but we have

$$K_1 = \phi_2 ,$$

$$K_2 = \phi_4/9 - \phi_2\phi_3 ,$$

$$K_3 = \phi_6/100 + (1 - 5\alpha_1)\phi_2\phi_5/10 + (3\alpha_1 - 2)\phi_3\phi_4/5 ,$$

and

$$K_4 = \Big\{3\phi_8 - 21\phi_2\phi_7/(1 - \alpha_1) + 9[35(1 - \alpha_3) - 3/(1 - \alpha_2)]\phi_3\phi_6 - 25(9 - 44\alpha_3 + 42\alpha_3^2)\phi_4\phi_5\Big\}/3675 ,$$

where

$$\phi_i = \frac{f^{(i)}(\zeta)}{i!\,f'(\zeta)} .$$

5. RELATED NONLINEAR RUNGE-KUTTA METHODS

The ordinary differential equation

(5.1) $$dx/dt = g(x) , \quad x(t_0) = x_0 ,$$

may be solved by quadrature and zero-finding: to find $x(t_0 + h)$ we need to find a zero of

$$f(x) = \int_{x_0}^{x} \frac{du}{g(u)} - h .$$

Note that $f(x_0) = -h$ is known, and $f'(x) = 1/g(x)$ may be evaluated almost as easily as $g(x)$. Thus, the zero-finding methods of Section 4 may be used to estimate $x(t_0 + h)$, then $x(t_0 + 2h)$, etc. When written in terms of g rather than f, the methods are seen to be similar to Runge-Kutta methods.

For example, the fourth-order zero-finding methods of Section 2 (with x_1 an exact zero of the quadratic $Q(x)$) gives:

$$g_0 = g(x_0) ,$$

$$\Delta = hg_0 ,$$

$$g_1 = g(x_0 + 2\Delta/3) ,$$

and

(5.2) $$x_1 = x_0 + 2\Delta/[1 + (3g_0/g_1 - 2)^{\frac{1}{2}}] .$$

Note that (5.1) is nonlinear in g_0 and g_1 , unlike the usual Runge-Kutta methods. (This makes it difficult to generalize our methods to systems of differential equations.) Since the zero-finding method is fourth-order, $x_1 = x(t_0 + h) + O(h^4)$, so our nonlinear Runge-Kutta method has order <u>three</u> by the usual definition of order (Henrici [62]).

Similarly, any of the zero-finding methods of Section 4 have a corresponding nonlinear Runge-Kutta method. Thus, we have:

<u>Theorem 5.1</u>

If $\nu > 0$, there is an explicit, nonlinear, Runge-Kutta method of order $2\nu - 1$, using ν evaluations of g per iteration, for single differential equations of the form (5.1).

By the result of Meersman and Wozniakowski, mentioned in Section 1, the order $2\nu - 1$ in Theorem 5.1 is the best possible. Butcher [65] has shown that the order of <u>linear</u> Runge-Kutta methods, using ν evaluations of g per iteration, is at most ν , which is less than the order of our methods if $\nu > 1$ (though the linear methods may also be used for systems of differential equations).

6. SOME NUMERICAL RESULTS

In this section we give some numerical results obtained with the nonlinear Runge-Kutta methods of Section 5. Consider the differential equation (5.1) with

(6.1) $$g(x) = (2\pi)^{\frac{1}{2}}\exp(x^2/2)$$

and $x(0) = 0$. Using step sizes $h = 0.1$ and 0.01, we estimated $x(0.4)$, obtaining a computed value x_h . The

error e_h was defined by

$$e_h = (2\pi)^{-\frac{1}{2}} \int_0^{x_h} \exp(-u^2/2)du - 0.4 .$$

All computations were performed on a Univac 1108 computer, with a floating-point fraction of 60 bits. The results are summarized in Table 6.1. The first three methods are derived from the zero-finding methods of Section 4 (with $\nu = 2, 3$ and 4 respectively). Method RK4 is the classical fourth-order Runge-Kutta method of Kutta [01], and method RK7 is a seventh-order method of Shanks [66].

<u>Table 6.1</u>: <u>Comparison of Runge-Kutta Methods</u>

Method	g evaluations per iteration	Order	$e_{0.1}$	$e_{0.01}$
Sec. 4	2	3	-9.45'-6	1.49'-7
Sec. 4	3	5	3.16'-6	-2.47'-11
Sec. 4	4	7	3.86'-8	3.69'-15
RK4	4	4	1.95'-5	7.90'-9
RK7	9	7	-5.19'-7	-1.67'-13

More extensive numerical results are given in Brent [75]. Note that the differential equation (6.1) was chosen only for illustrative purposes: there are several other ways of computing quantiles of the normal distribution. A practical application of our methods (computing quantiles of the incomplete Gamma and other distributions) is described in Brent [76].

7. OTHER ZERO-FINDING METHODS

In Section 1 we stated some generalizations of our methods (see Theorem 1.2). Further generalizations are described in Meersman [75]. Kacewicz [75] has considered methods which use information about an integral of f instead of a derivative of f .

"Sporadic" methods using derivatives may be derived as in Sections 2 and 3. For example, is there an eighth-order method which uses evaluations of f , f' , f'' , and f''' at x_0 , followed by evaluations of f' , f'' and f''' at some point y_1 ? Proceeding as in Sections 2 and 3 , we need a nonzero α satisfying

$$\text{rank}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 5\alpha & 6\alpha^2 & 7\alpha^3 \\ 12 & 20\alpha & 30\alpha^2 & 42\alpha^3 \\ 24 & 60\alpha & 120\alpha^2 & 210\alpha^3 \end{bmatrix} = 3 ,$$

which reduces to

$$35\alpha^3 - 84\alpha^2 + 70\alpha - 20 = 0 . \qquad (7.1)$$

Since (7.1) has one real root, $\alpha = 0.7449...$, an eighth-order method does exist. It is interesting to note that (7.1) is equivalent to the condition

$$\int_0^1 x^3(x - \alpha)^3 dx = 0 .$$

As a final example, we consider sixth-order methods using $f(x_0)$, $f'(x_0)$, $f''(y_1)$, and $f'''(y_2)$. (These could be called Abel-Gončarov methods.) Proceeding as above, we need α_1 and α_2 such that

$$\text{rank}\begin{bmatrix} 2 & 6\alpha_1 & 12\alpha_1^2 & 20\alpha_1^3 \\ 0 & 6 & 24\alpha_2 & 60\alpha_2^2 \\ 1 & 1 & 1 & 1 \end{bmatrix} = 2 ,$$

which gives

$$60\alpha_1^4 - 80\alpha_1^3 + 60\alpha_1^2 - 24\alpha_1 + 3 = 0 \qquad (7.2)$$

and

$$\alpha_2 = (1 - 6\alpha_1^2)/(4 - 12\alpha_1) .$$

Fortunately, (7.2) has two real roots, $\alpha_1 = 0.2074...$ and $\alpha_1 = 0.5351...$ Choosing one of these, we may evaluate $f(x_0)$, $f'(x_0)$ and $f''(y_1)$, where y_1 is defined as in Section 3. We may then fit a quadratic to the data, compute the perturbed $\tilde{\alpha}_1$, and take

$$\tilde{\alpha}_2 = (1 - 6\tilde{\alpha}_1^2)/(4 - 12\tilde{\alpha}_1) ,$$

etc., as in Section 3. It is not known whether this method can be generalized, i.e., whether real methods of order $2n$, using evaluations of $f(x_0)$, $f'(x_0)$, $f''(y_1)$, ..., $f^{(n)}(y_{n-1})$, exist for all positive n .

8. ACKNOWLEDGEMENT

The suggestions of J.C. Butcher, R. Meersman, M.R. Osborne and J.F. Traub are gratefully acknowledged.

REFERENCES

de Boor and Swartz [73] de Boor, C.W.R. and Swartz, B., "Collocation at Gaussian points", SIAM J. Numer. Anal. 10 (1973), 582-606.

Brent [75] Brent, R.P., "Some high-order zero-finding methods using almost orthogonal polynomials", J. Austral. Math. Soc. (Ser. B) 1 (1975). Also available as "Efficient methods for finding zeros of functions whose derivatives are easy to evaluate", Dept. of Computer Science, Carnegie-Mellon Univ. (Dec. 1974).

Brent [76] Brent, R.P. and E.M., "Efficient computation of inverse distribution functions", to appear.

Brent, Winograd and Wolfe [73] Brent, R.P., Winograd, S. and Wolfe, P., "Optimal iterative processes for root-finding", Numer. Math. 20 (1973), 327-341.

Butcher [65] Butcher, J.C., "On the attainable order of Runge-Kutta methods", Math. Comp. 10 (1965), 408-417.

Henrici [62] Henrici, P., "Discrete variable methods in ordinary differential equations", Wiley, New York, 1962.

Jarratt [69] Jarratt, P., "Some efficient fourth-order multipoint methods for solving equations", BIT 9 (1969), 119-124.

Jarratt [70] Jarratt, P., "A review of methods for solving nonlinear algebraic equations in one variable", in "Numerical methods for nonlinear algebraic equations", (edited by P. Rabinowitz), Gordon and Breach, New York, 1970, 1-26.

Kacewicz [75] Kacewicz, B., "An integral-interpolatory iterative method for the solution of nonlinear scalar equations", Dept. of Computer Science, Carnegie-Mellon Univ. (Jan. 1975).

Kung and Traub [73] Kung, H.T. and Traub, J.F., "Optimal order and efficiency for iterations with two evaluations", Dept. of Computer Science, Carnegie-Mellon Univ. (Nov. 1973). To appear in SIAM J. Num. Anal.

Kung and Traub [74] Kung, H.T. and Traub, J.F., "Optimal order of one-point and multipoint iteration", J. ACM 21 (1974), 643-651.

Kutta [01] Kutta, W., "Beitrag zur näherungsweisen Integration totaler Differentialgleichungen", Z. Math. Phys. 46 (1901), 435-452.

Meersman [75] Meersman, R., "Optimal use of information in certain iterative processes", these Proceedings.

Shanks [66] Shanks, E.B., "Solutions of differential equations by evaluations of functions", Math. Comp. 20 (1966), 21-38.

Traub [64] Traub, J.F., "Iterative methods for the solution of equations", Prentice-Hall, Englewood Cliffs, New Jersey (1964).

Wozniakowski [75a] Wozniakowski, H., "Properties of maximal order methods for the solution of nonlinear equations", ZAMM 55 (1975), 268-271.

Wozniakowski [75b] Wozniakowski, H., "Maximal order of multipoint iterations using n evaluations," these Proceedings.