Multipoint iterative methods for solving certain equations

By P. Jarratt*

A class of methods for solving equations is described which is very efficient in cases where the derivative can be rapidly evaluated compared with the function. Third- and fourth-order methods are analyzed and it is shown that a fifth-order method exists which requires only one function and three derivative evaluations per iteration.

1. Introduction

Multipoint iterative techniques for finding zeros of a function, $f(x)$, have been studied by Householder (1953), Ostrowski (1956) and Traub (1964). Briefly, these methods calculate new approximations to a zero of $f(x)$ by sampling, at each iteration, f and possibly its derivatives for a number of values of the independent variable. An additional feature of these processes is that they may possess a number of free parameters which can be used to ensure, for example, that the convergence is of a certain order for simple zeros, and that the sampling of the function and its derivatives is done at felicitous points. It is worth noting that the second condition is also a distinguishing characteristic of Gaussian quadrature formulae and of Runge–Kutta methods for integrating ordinary differential equations. However, although of considerable theoretical interest, multipoint iterative methods have not attracted any great practical attention, being usually rather inefficient computationally compared with more standard techniques. Nevertheless, for a commonly occurring class of functions for which the evaluation of the derivative f' is cheap compared with f, certain of these methods are undoubtedly attractive. In particular, Traub (Traub, 1964,—Chapter 9) has described third- and fourth-order formulae which require only one function and two or three derivative evaluations respectively per iteration. Thus for functions which are defined, for example, by integrals, a high-order root-finder is available which costs virtually no more per iteration than Newton's method or the *regula falsi*. In this paper, a class of multipoint methods is examined which yields a number of interesting third- and fourth-order processes, applicable to the same type of problem. It is further shown that a very economical fifth-order method can be constructed which costs only one function and three derivative evaluations per iteration.

2. Third-order formulae

In order to obtain solutions of $f(x) = 0$ (2.1)
we consider first a family of iterative methods defined by the formula

$$x_{n+1} = x_n - \frac{f(x_n)}{a_1\omega_1(x_n) + a_2\omega_2(x_n)} \tag{2.2}$$

* *Bradford Institute of Technology, Bradford, Yorks.*

where $\omega_1(x_n) = f'(x_n)$, $\omega_2(x_n) = f'[x_n + \alpha u(x_n)]$ and $u = f/f'$.

For non-zero values of the parameters a_1, a_2 and α, each application of (2.2) will require one evaluation of f and two of f'.

We now study the properties of the iteration (2.2) by assuming a simple root of (2.1) at $x = \theta$ and defining the error ϵ_n in the nth approximation by $x_n = \epsilon_n + \theta$.

Using the Taylor expansions of $f(x_n)$ and $f'(x_n)$ about the root $x = \theta$ we have $f(x_n) = \sum_{r=1}^{\infty} c_r\epsilon_n^r$ and

$$f'(x_n) = \sum_{r=1}^{\infty} rc_r\epsilon_n^{r-1} \text{ where } c_r = f^{(r)}(\theta)/r! \text{ and } c_0 = f(\theta) = 0.$$

From these results we find that

$$u(x_n) = f(x_n)/f'(x_n) = \epsilon_n - \frac{c_2}{c_1}\epsilon_n^2 + O[\epsilon_n^3]$$

and hence

$$\omega_2(x_n) = f'[x_n + \alpha u(x_n)] = c_1 + 2c_2(1+\alpha)\epsilon_n + \left[3c_3(1+\alpha)^2 - 2\frac{c_2^2}{c_1}\alpha\right]\epsilon_n^2 + O[\epsilon_n^3].$$

Thus

$$a_1\omega_1(x_n) + a_2\omega_2(x_n) = p_1 + p_2\epsilon_n + p_3\epsilon_n^2 + O[\epsilon_n^3]$$

where $p_1 = c_1(a_1 + a_2)$, $p_2 = 2c_2[a_1 + (1+\alpha)a_2]$

and
$$p_3 = 3c_3a_1 + a_2\left[3c_3(1+\alpha)^2 - 2\frac{c_2^2}{c_1}\alpha\right]. \tag{2.3}$$

Substituting now in (2.2) and expanding, we derive ultimately the relation

$$\epsilon_{n+1} = \left(1 - \frac{c_1}{p_1}\right)\epsilon_n + \frac{1}{p_1}\left(\frac{p_2}{p_1}c_1 - c_2\right)\epsilon_n^2 + \frac{1}{p_1}\left[\frac{p_2}{p_1}c_2 + \left(\frac{p_3}{p_1} - \frac{p_2^2}{p_1^2}\right)c_1 - c_3\right]\epsilon_n^3 + O[\epsilon_n^4]. \tag{2.4}$$

We can now use the free parameters a_1 and a_2 to ensure that the iteration (2.2) will be third-order for simple zeros of arbitrary functions f. For this we require

$$1 - \frac{c_1}{p_1} = 0$$

$$\frac{p_2}{p_1}c_1 - c_2 = 0$$

and using (2.3), these reduce to $a_1 + a_2 = 1$

$$2\alpha a_2 = -1,$$

the solution in terms of the remaining free parameter α being

$$a_1 = \frac{1+2\alpha}{2\alpha},\ a_2 = \frac{-1}{2\alpha}. \tag{2.5}$$

Furthermore, the asymptotic error constant, i.e. the coefficient of ϵ_n^3 in (2.4) now becomes

$$\frac{c_2^2}{c_1^2} - \frac{c_3}{c_1}\left(1 + \frac{3}{2}\alpha\right). \tag{2.6}$$

A family of third-order formulae can be obtained by assigning specific values to α. In deciding these values, we note from (2.2) that f' is sampled at x_n and $x_n + \alpha u(x_n)$. Hence by choosing α negative we shall ensure that in most cases of interest $x_n + \alpha u(x_n)$ will be nearer than x_n to θ. If $\alpha = -1$, then, of course, $x_n + \alpha u(x_n)$ is the point which would be predicted by Newton's formula. With these considerations in mind, two values of α are of special interest. First, if $\alpha = -\frac{1}{2}$, then from (2.5), $a_1 = 0$ and (2.2) is simplified to

$$x_{n+1} = x_n - \frac{f(x_n)}{f'[x_n - \frac{1}{2}u(x_n)]},$$

the asymptotic error constant having the value $\dfrac{c_2^2}{c_1^2} - \dfrac{c_3}{4c_1}$.

Secondly, from (2.6), if $\alpha = -2/3$ then the asymptotic error constant becomes $\dfrac{c_2^2}{c_1^2}$ and with $a_1 = 1/4$, $a_2 = 3/4$, (2.2) is

$$x_{n+1} = x_n - \frac{f(x_n)}{4\left(f'(x_n) + 3f'[x_n - \frac{2}{3}u(x_n)]\right)}.$$

This type of formula has particular merits in cases where f satisfies a simple second-order differential equation, when the form of the asymptotic error constant can be used to speed convergence.

3. Fourth-order formulae and a fifth-order process

In deriving fourth-order formulae, we increase the number of disposable parameters by using an iteration of the form

$$x_{n+1} = x_n - \frac{f(x_n)}{a_1\omega_1(x_n) + a_2\omega_2(x_n) + a_3\omega_3(x_n)} \tag{3.1}$$

where ω_1 and ω_2 are as defined in (2.2) and

$$\omega_3(x_n) = f'\left[x_n + \beta u(x_n) + \gamma\frac{f(x_n)}{\omega_2(x_n)}\right].$$

Each application of (3.1) will normally require one function and three derivative evaluations per iteration. The analysis of (3.1) now proceeds in a similar fashion to that of the third-order case.

We have

$$u(x_n) = \epsilon_n - \frac{c_2}{c_1}\epsilon_n^2 + 2\left(\frac{c_2^2}{c_1^2} - \frac{c_3}{c_1}\right)\epsilon_n^3 + O[\epsilon_n^4] \tag{3.2}$$

and hence

$$\begin{aligned}\omega_2(x_n) &= f'[x_n + \alpha u(x_n)]\\ &= c_1 + q_1\epsilon_n + q_2\epsilon_n^2 + q_3\epsilon_n^3 + O[\epsilon_n^4] \qquad (.\ 3)\end{aligned}$$

where $q_1 = 2c_2(1+\alpha)$, $q_2 = 3c_3(1+\alpha)^2 - 2\dfrac{c_2^2}{c_1}\alpha$

and $$q_3 = 4c_2\left(\frac{c_2^2}{c_1^2} - \frac{c_3}{c_1}\right)\alpha - 6\frac{c_2c_3}{c_1}\alpha(1+\alpha) + 4c_4(1+\alpha)^3.$$

We now find

$$\frac{f(x_n)}{\omega_2(x_n)} = \epsilon_n + r_1\epsilon_n^2 + r_2\epsilon_n^3 + O[\epsilon_n^4] \tag{3.4}$$

where $r_1 = -\dfrac{c_2}{c_1}(1+2\alpha)$

and $$r_2 = 2\frac{c_2^2}{c_1^2}(2\alpha^2 + 4\alpha + 1) - \frac{c_3}{c_1}(3\alpha^2 + 6\alpha + 2)$$

and substituting (3.2) and (3.4) into $\omega_3(x_n)$ we obtain

$$\begin{aligned}\omega_3(x_n) &= f'\left[x_n + \beta u(x_n) + \gamma\frac{f(x_n)}{\omega_2(x_n)}\right]\\ &= c_1 + s_1\epsilon_n + s_2\epsilon_n^2 + s_3\epsilon_n^3 + O[\epsilon_n^4] \qquad (3.5)\end{aligned}$$

where $$s_1 = 2c_2(1+\beta+\gamma),$$

$$s_2 = 2c_2\left(\gamma r_1 - \frac{c_2}{c_1}\beta\right) + 3c_3(1+\beta+\gamma)^2$$

and $$\begin{aligned}s_3 = 2c_2&\left[\gamma r_2 + 2\beta\left(\frac{c_2^2}{c_1^2} - \frac{c_3}{c_1}\right)\right]\\ &+ 6c_3(1+\beta+\gamma)\left(\gamma r_1 - \frac{c_2}{c_1}\beta\right)\\ &+ 4c_4(1+\beta+\gamma)^3.\end{aligned}$$

Finally, by substituting (3.3) and (3.5) into (3.1) and expanding we find, after some lengthy algebra, that the conditions for (3.1) to be fourth-order are

$$\left.\begin{aligned}a_1 + a_2 + a_3 &= 1\\ \alpha a_2 + (\beta+\gamma)a_3 &= -\tfrac{1}{2}\\ \alpha^2 a_2 + (\beta+\gamma)^2 a_3 &= \tfrac{1}{3}\\ \alpha\gamma a_3 &= \tfrac{1}{6}.\end{aligned}\right\} \tag{3.6}$$

Furthermore, using the relations of (3.6), the asymptotic error constant turns out to be of the form

$$2\frac{c_2^3}{c_1^3}(1+\alpha) - 3\frac{c_2c_3}{c_1^2}\left[1 + \frac{1}{2}\{\alpha + 2(\beta+\gamma)\}\right] + \frac{c_4}{c_1}\left[1 + \frac{4}{3}(\alpha+\beta+\gamma) + 2\alpha(\beta+\gamma)\right]. \tag{3.7}$$

(3.6) represents a set of four equations in six unknowns and hence we may solve for any four of the variables in terms of the remaining two. However, the value of α decides where $\omega_2(x_n)$ is sampled, and from (3.5) we see that $\omega_3(x_n)$ is sampled at a point determined effectively by the value of $\beta + \gamma$. We therefore write $\theta = \beta + \gamma$ and solve (3.6) for a_1, a_2, β and γ in terms of θ and α.

For $\alpha \neq -\frac{2}{3}$, $\theta \neq 0$, $\theta \neq \alpha$, the solution is

$$a_1 = \frac{6\alpha\theta + 3(\alpha+\theta) + 2}{6\alpha\theta},\ a_2 = \frac{3\theta+2}{6\alpha(\alpha-\theta)},\ a_3 = \frac{3\alpha+2}{6\theta(\theta-\alpha)},$$

$$\beta = \theta - \frac{3\theta(\theta-\alpha)}{2\alpha(3\alpha+2)} \text{ and } \gamma = \frac{3\theta(\theta-\alpha)}{2\alpha(3\alpha+2)}.$$

The last equation of (3.6) shows that we cannot have α, γ or a_3 zero. We dispose of the solutions which are special cases first. $\theta = 0$ is immediately excluded from consideration on sampling grounds. Also $\alpha = -\frac{2}{3}$ implies $\theta = \alpha$, and we can now find a one-parameter family of solutions of the form

$$a_1 = \frac{1}{4},\ a_2 = \frac{3}{4} + \frac{3}{8\gamma},\ a_3 = -\frac{3}{8\gamma},$$

$$\alpha = -\frac{2}{3} \text{ and } \beta = -\frac{2}{3} - \gamma,\ \gamma = -\frac{1}{2}$$

giving one solution of interest.

Assuming henceforth that $\alpha \neq \theta$, we go on to seek solutions which simplify the form of (3.7), the asymptotic error constant. Now the second and third terms of (3.7) will vanish for values of α and θ satisfying the simultaneous equations

$$\alpha + 2\theta = -2$$

$$4(\alpha + \theta) + 6\alpha\theta = -3,$$

which have solutions $\alpha = -\frac{1}{3}$, $\theta = -\frac{5}{6}$ and $\alpha = -1$, $\theta = -\frac{1}{2}$. We see that the second solution is of particular interest since $\alpha = -1$ also renders the first term of (3.7) zero, and we obtain a formula which is fifth-order. The corresponding values of the parameters are $a_1 = \frac{1}{6}$, $a_2 = \frac{1}{6}$, $a_3 = \frac{2}{3}$, $\beta = -\frac{1}{8}$ and $\gamma = -\frac{3}{8}$. The asymptotic error constant of this fifth-order process has not been estimated. The solution $\alpha = -\frac{1}{3}$, $\theta = -\frac{5}{6}$ results in a fourth-order process whose asymptotic error constant depends only on c_2 and c_1, the values of the parameters being $a_1 = \frac{1}{10}$, $a_2 = \frac{1}{2}$, $a_3 = \frac{2}{5}$, $\alpha = -\frac{1}{3}$, $\beta = \frac{25}{24}$ and $\gamma = -\frac{15}{8}$. Such processes are valuable since, as Traub (Ch. 11) has shown, they may be readily generalized to deal with the problem of solving systems of equations. Other fourth-order formulae having useful properties can also be constructed, for example, with $a_1 = 0$ or $a_2 = 0$, but the details will be omitted here.

4. Multiple roots

In the case where the root θ of (2.1) is multiple, the convergence of the iterative formulae derived from (2.2) and (3.1) in general falls to first-order. Thus for the iteration (2.2) we find, by setting $c_1 = 0$ and expanding, that for a double root, the errors are related by

$$\epsilon_{n+1} = \left[1 - \frac{1}{2\left[a_1 + a_2\left(1 + \frac{\alpha}{2}\right)\right]}\right]\epsilon_n + O[\epsilon_n^2],$$

and using (2.5) this reduces to $\epsilon_{n+1} = \frac{1}{3}\epsilon_n + O[\epsilon_n^2]$.

This means that asymptotically only one extra significant figure approximately is gained every three iterations. In practice one method of dealing with multiple roots is to compute at each iteration the quantity

$$R_n = \left|\frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}}\right|.$$

Using $x_n = \epsilon_n + \theta$, we have

$$R_n = \left|\frac{\epsilon_n - \epsilon_{n-1}}{\epsilon_{n-1} - \epsilon_{n-2}}\right|$$

and assuming, for multiple roots, that $\epsilon_{n+1} \sim K\epsilon_n$ where K is a constant, it follows that $R_n \sim |K|$. Thus, if during a calculation R_n has remained approximately constant over a number of iterations, then it is safe to assume geometric convergence and apply an accelerating device, for example Aitken's δ^2 process. Another possibility is to carry out the calculation replacing f by f/f' which has only simple zeros, but this technique is rather wasteful since in practice the roots of interest are usually strongly isolated.

5. Conclusions

The formulae derived in Sections 3 and 4 will provide very efficient methods for locating simple zeros of functions whose derivatives can be rapidly computed compared with the function itself. A limited amount of numerical experience has indicated that a reasonably close starting value x_0 is necessary for the methods to converge; this condition, however, applies to practically all iterative methods for solving equations.

References

HOUSEHOLDER, A. S. (1953). *Principles of Numerical Analysis*, New York, McGraw-Hill, pp. 126-128.

OSTROWSKI, A. M. (1956). "Über Verfahren von Steffenson und Householder zur Konvergenzverbesserung von Iterationen." *Z. Angew. Math. Phy* ., Vol. 7, pp. 218-219.

TRAUB, J. F. (1964). *Iterative Methods for the solution of Equations*, Englewood Cliffs, N. J. Prentice-Hall.