NOTES ON ITERATIVE PROCESSES

By D. R. HARTREE

Received 10 May 1948

1. *Introduction.* By an 'iterative process' for the solution of an equation

$$F(y) = 0 \tag{1}$$

is meant here a method by which we attempt to solve the equation by writing it in the form

$$y = f(y) \tag{2}$$

and constructing a sequence $\{y_n\}$ such that

$$y_{n+1} = f(y_n). \tag{3}$$

If this sequence converges to a limit Y, then $y = Y$ is a solution of equation (1). Here y may be a number, a vector, or a matrix. It may also be a function of one or more continuous variables x; in this case the equation to be solved will probably involve differential, integral, or finite difference operators, and $F(y)$, $f(y)$ must be understood in a correspondingly generalized sense. For the main part of the present note it will be assumed that y is a number.

Iterative methods seem very suitable, in principle, for application in large automatic calculating machines such as the ENIAC (1) and others under construction or planned, on account of the self-checking and almost automatic character of such methods, and it seems likely that the development of such machines will stimulate developments in the planning and use of iterative methods. Two developments required are mentioned in § 8.

An important feature of iterative methods is the character of the variation of the error of the result with the number n of repetitions of the iterative process. This has probably been recognized by most workers who have had practical experience of using such methods; but I know of no formal treatment, hence the present note.

2. *The order of an iterative process.* If Y is a solution of $F(y) = 0$, so that $Y = f(Y)$, let

$$y_n = Y + \eta_n, \tag{4}$$

so that η_n is the error of the nth iterate. Now let the right-hand side of (3) be expanded in a Taylor series about $y = Y$; then (3) becomes

$$\eta_{n+1} = a_1\eta_n + a_2\eta_n^2 + a_3\eta_n^3 + \ldots, \tag{5}$$

where

$$a_1 = F'(Y), \quad a_2 = \tfrac{1}{2}F''(Y), \quad a_3 = \frac{1}{3!}F'''(Y)\ldots,$$

and the characteristic feature mentioned depends on the smallest value of k for which a_k [or $F^{(k)}(Y)$] is non-zero. If this $k = K$, the iterative process will be described as 'Kth order'. First- and second-order processes are common in practical numerical analysis; third-order processes are uncommon, and I do not know any examples of higher-order process having been used..

https://doi.org/10.1017/S0305004100024762 Published online by Cambridge University Press

For a first-order process ($K = 1$), $a_1 \neq 0$, and then, for sufficiently small η,

$$\eta_{n+1} = a_1 \eta_n, \tag{6}$$

so that, if this approximation is already adequate for $n = 0$,

$$\eta_n = a_1^n \eta_0. \tag{7}$$

Thus in this case the errors of successive iterates are in geometrical progression, and the number of additional correct figures obtained for each repeat of the iterative process (or more often in practice, the number of repeats required to obtain each new correct figure) is independent of n.

For a second-order process, $a_1 = 0$, $a_2 \neq 0$, so that the leading term in (5) is

$$\eta_{n+1} = a_2 \eta_n^2, \tag{8}$$

and so

$$a_2 \eta_{n+1} = (a_2 \eta_0)^{2^n}.$$

In this case the number of correct figures is approximately doubled by each repeat of the iterative process, so that such a process is very much to be preferred to a first-order process, at least once a fair approximation to Y has been reached. Similarly for a third-order process

$$\eta_{n+1} = a_3 \eta_n^3, \quad a_3^{\frac{1}{2}} \eta_n = (a_3^{\frac{1}{2}} \eta_0)^{3^n}, \tag{9}$$

and the number of correct figures is approximately trebled by each repeat of the iterative process.

3. *Non-uniqueness of iterative processes of given order for given equation.* For practical purposes one wants to have a process which is at least second order, and on trying to devise such a process the question arises whether, for a given equation $F(y) = 0$, there is a *unique* second-order process or several among which we might be able to select one as the most convenient for practical work. Or, in general, if there is a Kth order process, is it unique?

The following argument shows that if there is a $(K+1)$th-order process, the Kth order is not unique. Suppose $y = f_{\mathrm{I}}(y)$ gives a Kth-order process and $y = f_{\mathrm{II}}(y)$ gives a $(K+1)$th-order process. Then for any constant q, $y = qf_{\mathrm{I}}(y) + (1-q) f_{\mathrm{II}}(y)$ also gives a process of at least Kth order. And conversely it is possible from two Kth-order processes to derive a $(K+1)$th-order process, provided the values of a_{K+1} in the two Kth-order processes are unequal. It will be shown later how, for some simple equations, iterative processes of any order can be derived, so that in these cases there is not a unique process of any order, and it seems likely that this is general.†

As an example of non-uniqueness of a second-order process, consider the question of iterative processes for a square root. The best known one is Newton's, given by

$$f_{\mathrm{I}}(y) = \tfrac{1}{2}(y + a/y) = \tfrac{1}{2}(y^2 + a)/y, \tag{10}$$

for which

$$Y = a^{\frac{1}{2}}, \quad f_{\mathrm{I}}'(y) = \tfrac{1}{2}(y^2 - a)/y^2, \quad f_{\mathrm{I}}'(Y) = 0, \quad f_{\mathrm{I}}''(Y) = 1/Y. \tag{11}$$

Another, derivable from a second-order process for an inverse square root [see (19) below], is

$$f_{\mathrm{II}}(y) = 2y^3/(3y^2 - a), \tag{12}$$

for which

$$Y = a^{\frac{1}{2}}, \quad f_{\mathrm{II}}'(y) = 6y^2(y^2 - a)/(3y^2 - a)^2, \quad f_{\mathrm{II}}'(Y) = 0, \quad f_{\mathrm{II}}''(Y) = 3/Y. \tag{13}$$

† This conjecture has been confirmed by C. Domb (see the following paper, p. 237).

https://doi.org/10.1017/S0305004100024762 Published online by Cambridge University Press

Another, see (28) below, is $$f_{\mathrm{III}}(y) = y(3a-y^2)/2a, \tag{14}$$

for which $$Y = a^{\frac{1}{2}}, \quad f'_{\mathrm{III}}(y) = 3(a-y^2)/2a, \quad f'_{\mathrm{III}}(Y) = 0, \quad f''_{\mathrm{III}}(Y) = -3/Y. \tag{15}$$

Formulae (10), (12) and (14) exhibit three distinct second-order iterative processes for $a^{\frac{1}{2}}$. By combining any two of them, a third-order process can be derived; for example, if $$f_{\mathrm{IV}}(y) = \tfrac{3}{2}f_{\mathrm{I}}(y) - \tfrac{1}{2}f_{\mathrm{II}}(y), \tag{16}$$

$f_{\mathrm{IV}}(Y) = Y$, and it follows from (11) and (13) that $f'_{\mathrm{IV}}(Y) = 0$, $f''_{\mathrm{IV}}(Y) = 0$. So the rather unpromising-looking function

$$f_{\mathrm{IV}}(y) = (5y^4 + 6ay^2 - 3a^2)/4y(3y^2 - a) \tag{17}$$

actually gives a third-order process for $a^{\frac{1}{2}}$. For $a = 10$, $y_0 = 3$, it gives $y_1 = 3{\cdot}161765$, which is already correct to four figures. In practice it would probably be more convenient to evaluate $f_{\mathrm{IV}}(y)$ from (16) than from (17).

Similarly, by making a linear combination of all three of (10), (12) and (14), a fourth-order process can be derived.

4. *Examples of iterative processes.* The following are examples of second-order processes: $$y_{n+1} = y_n(2 - zy_n), \quad Y = 1/z \tag{18}$$

for a reciprocal, $$y_{n+1} = y_n(3 - zy_n^2)/2, \quad Y = 1/z^{\frac{1}{2}} \tag{19}$$

for an inverse square root, and in general

$$y_{n+1} = y_n[(p+1) - zy_n^p]/p, \quad Y = 1/z^{1/p} \tag{20}$$

for an inverse pth root (Aiken (2)). The formula

$$y_{n+1} = y_n[(q+1) - zy_n^p]/q$$

for any q also giving an iterative process for $1/z^{1/p}$, but it is easily shown, by evaluating $f'(y)$, that unless $q = p$ this is only first order. Use of these formulae does not involve any division, and, as has already been explained, this may be an advantage in using automatic machines; p being known, $1/p$ can be calculated in advance and supplied to the machine as a known constant by which to multiply in evaluating the right-hand side of (20).

The Newton-Raphson process

$$y_{n+1} = y_n - \phi(y_n)/\phi'(y_n) \tag{21}$$

for a root of $\phi(y) = 0$ is also second order, since for

$$f(y) = y - \phi(y)/\phi'(y), \quad f'(y) = \phi(y)\phi''(y)/\{\phi'(y)\}^2, \quad f'(Y) = 0.$$

Formulae (18)–(20) can be extended as follows. Let $f(y) = y\psi(y)$; then the conditions $f(y) = Y, f'(Y) = 0, f''(Y) = 0, \ldots$ are

$$\left.\begin{aligned} \psi(Y) &= 1, \\ Y\psi'(Y) + \psi(Y) &= 0 \quad \text{or} \quad Y\psi'(Y) = -1, \\ Y\psi''(Y) + 2\psi'(Y) &= 0 \quad \text{or} \quad Y^2\psi''(Y) = +2, \\ Y\psi'''(Y) + 3\psi''(Y) &= 0 \quad \text{or} \quad Y^3\psi'''(Y) = -6, \\ &\text{etc.} \end{aligned}\right\} \tag{22}$$

https://doi.org/10.1017/S0305004100024762 Published online by Cambridge University Press

Now in trying to obtain an iterative process for $1/z^{1/p}$, let

$$\psi(y) = 1+\beta_1(1-zy^p)+\beta_2(1-zy^p)^2+\ldots, \tag{23}$$

so that

$$Y\psi'(Y) = -\beta_1 pzY^p = -\beta_1 p, \quad Y^2\psi''(Y) = -\beta_1 p(p-1)+2\beta_2 p^2, \quad \text{etc.}, \tag{24}$$

so that the second of equations (22) gives $\beta_1 = 1/p$ and the third gives $\beta_2 = (p+1)/2p^2$, etc. If now in (23) we put $\beta_1 = 1/p$, $\beta_k = 0$ for $k \geqslant 2$, we satisfy the first of formulae (22) but not the second, and consequently have a second-order process, namely, that given by formula (20). If we put $\beta_1 = 1/p$, $\beta_2 = (p+1)/2p^2$, $\beta_k = 0$ for $k \geqslant 3$, we satisfy the first two of formulae (22) and get a third-order process, given by

$$\begin{aligned} f(y) &= y[1+(1-zy^p)/p+(p+1)(1-zy^p)^2/2p^2] \\ &= y[2p^2+3p+1-2(2p+1)zy^p+(p+1)(zy^p)^2]/2p^2. \end{aligned} \tag{25}$$

It is possible to extend this argument and obtain higher-order processes, but the formulae get more complicated, and it seems probable that for practical purposes it would be more important to improve the convergence of the early stages of the iterative process rather than to improve further that of the later stages.

For positive powers, $a^{1/p}$, one can similarly put

$$f(y) = y[1-\beta_1(y^p-a)+\beta_2(y^p-a)^2\ldots]. \tag{26}$$

Then $$Y = a^{1/p}, \quad Y\psi'(Y) = -p\beta_1 Y^p = -p\beta_1 a,$$

so that a second-order process for $a^{1/p}$ is given by $\beta_1 = 1/pa$, that is

$$f(y) = y[(p+1)a-y^p]/pa. \tag{27}$$

In particular, for $p = 2$, $$f(y) = y(3a-y^2)/2a. \tag{28}$$

For practical purposes, this might be a good deal more convenient than the more familiar formula (10), since with some machines, the ENIAC for example, division is a much longer process than multiplication, and one to be avoided if possible, and in using (28) only one operation of division is required, to give $1/a$, whereas in using (10) or (12) a new division is required for each repetition of the iterative process. As an example of the use of (24), for $a = 10$, $y_0 = 3$ it gives

$$y_0 = 3, \quad y_1 = 3{\cdot}15, \quad y_2 = 3{\cdot}16220625, \quad y_3 = 3{\cdot}1622776577;$$

to 11 figures, $a^{\frac{1}{2}} = 3{\cdot}1622776602$, so that y_3 is correct to 9 figures.

5. *Derivation of a second-order process from a first-order process.* For equations for which the solution is a number, it is always possible to derive a second-order process from a first-order process as follows. Let y_0, y_1, y_2 be three successive iterates in a first-order process. If the errors η_n were exactly in geometrical progression, the value Y^* of y satisfying the equation would be such that

$$\frac{Y^*-y_2}{Y^*-y_1} = \frac{Y^*-y_1}{Y^*-y_0},$$

or

$$Y^* = (y_2y_0-y_1^2)/(y_2-2y_0+y_0) \tag{29}$$

$$= Y+(\eta_2\eta_0-\eta_1^2)/(\eta_2-2\eta_1-\eta_0). \tag{30}$$

https://doi.org/10.1017/S0305004100024762 Published online by Cambridge University Press

Since the iterative process by which y_{n+1} is derived from y_n is first order, equation (5) applies with $a_1 \neq 0$, so that

$$\eta_1 = a_1\eta_0 + a_2\eta_0^2 + O(\eta_0^3),$$

$$\eta_2 = a_1\eta_1 + a_2\eta_1^2 + O(\eta_1^3)$$

$$= a_1^2\eta_0 + a_1a_2(1+a_1)\eta_0^2 + O(\eta_0^3),$$

and hence from (30)

$$Y^* = Y + \frac{a_1a_2}{a_1-1}\eta_0^2 + O(\eta_0^3). \tag{31}$$

Let this value of Y^* be written Y_1^*. If now Y_1^*, instead of y_2, is taken as the starting point y_0 for another two repetitions of the first-order iterative process, leading to another quantity Y_2^* on application of (29), and so on, the process giving the successive values of Y_m^* is second order. This process was given by Aitken (3) and has also been considered more recently by Samuelson (4).

As it stands, formula (29) is not convenient for practical work, since if $y_0 = Y$, so that $y_1 = y_2 = Y$ also, the right-hand side of (29) takes the indeterminate form 0/0; and if y_0, y_1, y_2 are nearly equal, Y is badly determined by (29); we appear to have the paradoxical situation that the better the approximation y_0 is to the required solution of the equation, the less well determined Y^* is as a better approximation! This is a consequence of the fact that (29) expresses Y as the ratio of two small quantities each of order η_0. This feature of formula (29) is pointed out by Samuelson (4), but he only suggests care in dropping figures as a way of avoiding the difficulty. If, however, this formula is written so that Y^* is expressed as the sum of the best approximation, y_2, and a correction term, thus:

$$Y^* = y_2 - \frac{(y_1-y_2)^2}{y_2-2y_1+y_0}, \tag{32}$$

the numerator of this correction term is $O(\eta_0^2)$ whereas the denominator is $O(\eta_0)$, so the whole is $O(\eta_0)$, and there is no difficulty in evaluating it.

6. *Iterative processes for vectors and matrices.* If y is a vector, with components y_μ, then in place of (5) we have

$$(\eta_\mu)_{n+1} = a_{\mu\nu}(\eta_\nu)_n + O(\eta_\nu^2),$$

where

$$a_{\mu\nu} = \left(\frac{\partial F_\mu}{\partial y_\nu}\right)_{y=Y}, \tag{33}$$

and the summation convention is assumed. For a second-order process, the matrix $a_{\mu\nu}$ must be zero.

It may be expected that it will be more difficult to derive an iterative process in which the whole matrix $a_{\mu\nu}$ is zero than one in which one coefficient, a_1 in (5), is zero, so that second-order iterative processes for vector quantities are likely to be rarer than for scalar quantities. However, some simple examples of second-order processes can be given.

In formula (18) it is not necessary that y and z should be real; if they are complex, say $y = u+iv$, $z = a+ib$, separation of real and imaginary points gives

$$u_{n+1} = 2u_n - a(u_u^2+v_n^2) + 2bu_nv_n, \quad v_{n+1} = 2v_n - 2au_nv_n - b(u_n^2+v_n^2). \tag{34}$$

https://doi.org/10.1017/S0305004100024762 Published online by Cambridge University Press

It can easily be verified that all four components of the matrix (33) are zero, so that (34) does give a second-order process for

$$y = (u, v) = \left(\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2}\right).$$

The derivation of a second-order process from a first-order process by an analogue of the method of the previous section would seem to require the determination of the latent roots of the matrix (33) and it is doubtful if this would be practicable.

Some of the iterative process for numerical solutions can also be applied to equations whose solutions are matrices; in particular (18) can be used for inverting matrices, as pointed out by von Neumann.

7. *Iterative methods in the solution of differential equations.* In application to ordinary differential equations, iterative methods may be particularly useful in the solution of equations with two-point boundary conditions, in which they may avoid the evaluation of a number of trial solutions which diverge rapidly from one another as the integration proceeds.

For example, for the solution of

$$y''' + (1+y^2)\,y'' = \upsilon \tag{35}$$

with boundary conditions

$$y = y' = 0 \text{ at } x=0, \quad y' = 1 \text{ at } x = \infty,$$

we may form the sequence of functions y_n such that

$$y'''_{n+1} + (1+y_n^2)\,y''_{n+1} = 0; \tag{36}$$

the solution of this, satisfying all boundary conditions, is

$$y_{n+1} = \int_0 \frac{\int_0 \exp\left\{-\int_0 (1+y_n^2)\,dx\right\} dx}{\int_0^\infty \exp\left\{-\int_0 (1+y_n^2)\,dx\right\} dx} dx,$$

and by use of this the solution of (35) is reduced to a process of iterative quadrature. If we write

$$y_n(x) = y(x) + \eta_n(x),$$

then it follows from (35) and (36) that

$$\eta'''_{n+1} + (1+y^2)\,\eta''_{n+1} + 2yy''\eta_n = O(\eta_n^2),$$

so that η_{n+1} is $O(\eta_n)$, and the process is first order.

A similar, but more involved, example is provided by some equations arising in the theory of the laminar boundary layer in a compressible fluid (5).

8. *Developments required.* Part of the purpose of this paper is to call attention to two aspects of iterative methods in which developments are required. The first is a general method of obtaining iteration formulae of second and higher orders for the solution of any algebraic equation ('algebraic' here is meant as an antithesis to 'differential'—in this sense the equation $e^x = x^2$ is 'algebraic'); the Newton-Raphson formula is always available, but it is probably not unique, and if it is not, there may in any

https://doi.org/10.1017/S0305004100024762 Published online by Cambridge University Press

particular case be another method which is more convenient for practical use, in the way that formula (28) may be more convenient than (10) for square roots.

The second is the development of a practical method of deriving a second-order process from a first-order process in the case of iterative solution of a differential or integral equation. Such a development would be a real advance in the theory of iterative methods, and would probably find wide use in the application of electronic digital machines to such equations.

9. *Summary.* If η_n is the error in the result of repeating an iterative process n times, and η_{n+1} is $O(\eta_n^K)$, the iterative process is called Kth order. It is shown that if for a given equation there is an iterative process of the $(K+1)$th order, the iterative process of the Kth order is not unique, and conversely if an iterative process of the Kth order is not unique, it is generally possible to construct from two of the Kth order processes a process of the $(K+1)$th order. As an example, three second-order processes for a square root are exhibited, and a third-order process is derived from two of them.

Iterative processes for positive and negative integer roots are given, of kinds suitable for use on machines in which division is a relatively slow process and one to be used sparingly. It is shown how a second-order process can be derived from the results of two repetitions of a first-order process. The extension of this iterative process for the solution of differential and integral equations is a development which is urgently required.

REFERENCES

(1) See HARTREE, D. R. *Nature, London*, 158 (1946), 500.
(2) AIKEN, H. H. and others. *Manual of Operation of the Automatic Sequence-controlled Calculator* (Harvard University Press, 1946), 176.
(3) AITKEN, A. C. *Proc. Roy. Soc. Edinburgh*, 46 (1925), 289 (see § 8) and 57 (1936), 269 (see § 8).
(4) SAMUELSON, P. A. *J. Math. Phys.* 24 (1945), 131.
(5) COPE, W. F. and HARTREE, D. R. *Philos. Trans.* A, 241 (1948), 1 (see §§ 12, 13).

CAVENDISH LABORATORY
CAMBRIDGE

https://doi.org/10.1017/S0305004100024762 Published online by Cambridge University Press