University of Maryland College Park

Institute for Advanced Computer Studies TR–93–10
Department of Computer Science TR–3030

On the Convergence of Multipoint Iterations*

G. W. Stewart†

February, 1993
Reviseed, Janurary 1994

ABSTRACT

This note gives a new convergence proof for iterations based on multipoint formulas. It rests on the very general assumption that if the desired fixed point appears as an argument in the formula then the the formula returns the fixed point.

*This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports`.

†Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742.

On the Convergence of Multipoint Iterations

G. W. Stewart

ABSTRACT

This note gives a new convergence proof for iterations based on multipoint formulas. It rests on the very general assumption that if the desired fixed point appears as an argument in the formula then the the formula returns the fixed point.

Many useful iterations for finding a zero x_* of a function f generate iterates according to a scheme of the form

$$x_{k+1} = \varphi(x_k, x_{k-1}, \ldots, x_{k-n+1}). \tag{1}$$

For example, the secant method proceeds according to the formula

$$x_{k+1} = \frac{x_{k-1}f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}. \tag{2}$$

Since the method uses two previous iterates to generate the next, it is called a two-point method. Three-point methods are also in common use; e.g., Muller's method [3] and the method of inverse quadratic interpolation [2]. More generally, the iteration (1) is called an n-point iteration or generically a multipoint iteration.

Proofs of local convergence of multipoint methods usually involve using the special form of the iteration function φ to show that the errors $e_k = x_k - x_*$ satisfy an inequality of the form

$$|e_{k+1}| \leq C|e_k e_{k-1} \cdots e_{k-n+1}|, \tag{3}$$

from which it can be shown that the convergence is at least superlinear of order p, where p is the positive root of the equation

$$p^n - p^{n-1} - \cdots - p - 1 = 0. \tag{4}$$

Only the constant in (3) depends on the particular method; in other words, the order of convergence depends only on the number of points.

There is a simple reason why these different methods have the same rate of convergence: namely, each of the iterative methods named above return the solution whenever one of the arguments is the solution. More precisely, in the case of the secant method

$$\varphi(u, x_*) \equiv x_* \quad \text{and} \quad \varphi(x_*, v) \equiv x_*,$$

as can easily be seen from (2). The purpose of this note is to show that a general convergence proof for multipoint iterations can be based on this property.

Throughout we will assume that φ has as many derivatives as is needed for the analysis. The ideas are sufficiently well illustrated for the general two-point iteration, and to avoid clutter we will analyze that case. At the end of the note, we will indicate the modifications necessary for the general case.

The first step in the analysis is to observe that since $\varphi(u, x_*)$ and $\varphi(x_*, v)$ are constant their derivatives with respect to u and v are zero:

$$\varphi_u(u, x_*) \equiv 0 \quad \text{and} \quad \varphi_v(x_*, v) \equiv 0.$$

The same is true of the unmixed second derivatives:

$$\varphi_{uu}(u, x_*) \equiv 0 \quad \text{and} \quad \varphi_{vv}(x_*, v) \equiv 0.$$

The next step is to derive an error recursion. We begin by expanding φ about (x_*, x_*) in a Taylor series. Specifically,

$$\begin{aligned}\varphi(x_* + p, x_* + q) = x_* &+ \varphi_u(x_*, x_*)p + \varphi_v(x_*, x_*)q \\ &+ \frac{1}{2}[\varphi_{uu}(x_* + \theta p, x_* + \theta q)p^2 \\ &\quad + 2\varphi_{uv}(x_* + \theta p, x_* + \theta q)pq + \varphi_{vv}(x_* + \theta p, x_* + \theta q)q^2],\end{aligned}$$

where $\theta \in [0, 1]$. Since $\varphi_u(x_*, x_*) = \varphi_v(x_*, x_*) = 0$,

$$\begin{aligned}\varphi(x_* + p, x_* + q) = x_* &+ \frac{1}{2}[\varphi_{uu}(x_* + \theta p, x_* + \theta q)p^2 \\ &+ 2\varphi_{uv}(x_* + \theta p, x_* + \theta q)pq + \varphi_{vv}(x_* + \theta p, x_* + \theta q)q^2],\end{aligned} \tag{5}$$

We would like to factor the product pq out of this expression to obtain a product of errors like (3); however, we must first massage the terms in p^2 and q^2. Since $\varphi_{uu}(x_* + \theta p, x_*) = 0$, it follows from a Taylor expansion in the second argument that

$$\varphi_{uu}(x_* + \theta p, x_* + \theta q) = \varphi_{uuv}(x_* + \theta p, x_* + \tau_q \theta q)\theta q,$$

where $\tau_q \in [0,1]$. Similarly,

$$\varphi_{vv}(x_* + \theta p, x_* + \theta q) = \varphi_{uvv}(x_* + \tau_p \theta p, x_* + \theta q)\theta p,$$

where $\tau_p \in [0,1]$. Substituting these values in (5) gives

$$\begin{aligned}\varphi(x_* + p, x_* + q) = x_* + \frac{pq}{2}[&\varphi_{uuv}(x_* + \theta p, x_* + \tau_q \theta q)\theta p \\ &+ 2\varphi_{uv}(x_* + \theta p, x_* + \theta q) + \varphi_{uvv}(x_* + \tau_p \theta p, x_* + \theta q)\theta q].\end{aligned} \tag{6}$$

Turning now to the iteration itself, let the starting values be x_0 and x_1, and let their errors be $e_0 = x_0 - x_*$ and $e_1 = x_1 - x_*$. Taking $p = e_1$ and $q = e_0$ in (6), we get

$$\begin{aligned}e_2 &= \varphi(x_* + e_1, x_* + e_0) \\ &= x_* + \frac{e_1 e_0}{2}[\varphi_{uuv}(x_* + \theta e_1, x_* + \tau_{e_0}\theta e_0)\theta e_1 \\ &\qquad + \varphi_{uv}(x_* + \theta e_1, x_* + \theta e_0) + \varphi_{uvv}(x_* + \tau_{e_1}\theta e_1, x_* + \theta e_0)\theta e_0] \\ &\equiv \frac{e_1 e_0}{2} r(e_1, e_0).\end{aligned} \tag{7}$$

This is the error recurrence we need.

We are now ready to establish the convergence of the method. First note that

$$r(0,0) = 2\varphi_{uv}(x_*, x_*).$$

Hence there is a $\delta > 0$ such that if $|u|, |v| \le \delta$ then

$$|vr(u,v)| \le C < 1.$$

Now let $|e_0|, |e_1| \le \delta$. From the error recurrence (7) it follows that $|e_2| \le C|e_1| < |e_1| \le \delta$. Hence

$$|e_1 r(e_2, e_1)| \le C < 1.$$

and $|e_3| \le C|e_2| \le C^2|e_1|$. By induction

$$|e_k| \le C^{k-1}|e_1|,$$

and since the right-hand side of this inequality converges to zero, we have $e_k \to 0$; i.e., the general two-point iteration converges from any two starting values whose errors are less than δ in absolute value.

We now turn to the convergence rate of two point methods. The first thing to note is that since

$$e_{k+1} = \frac{e_k e_{k-1}}{2} r(e_k, e_{k-1}) \tag{8}$$

and $r(0,0) = 2\varphi_{uv}(x_*, x_*)$, we have

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k e_{k-1}} = \varphi_{uv}(x_*, x_*). \tag{9}$$

If $\varphi_{uv}(x_*, x_*) \neq 0$, we shall say that the sequence $\{x_k\}$ exhibits *two-point convergence.*

We are going to show that two-point convergence is superlinear of order

$$p = \frac{1+\sqrt{5}}{2} = 1.618\ldots.$$

Here p is the largest root of the equation

$$p^2 - p - 1 = 0. \tag{10}$$

There are two ways to establish this fact. The first is to derive (10) directly from (9), which is the usual approach. However, since we already know the value of p, we can instead set

$$s_k = \frac{|e_{k+1}|}{|e_k|^p} \tag{11}$$

and use (9) to verify that the s_k have a nonzero limit, which is usual definition of pth order convergence.

From (11) we have

$$|e_k| = s_{k-1}|e_{k-1}|^p$$

and

$$|e_{k+1}| = s_k|e_k^p| = s_k s_{k-1}^p |e_{k-1}|^{p^2}.$$

From (8),

$$|r_k| \equiv |r(e_k, e_{k-1})| = \frac{s_k s_{k-1}^p |e_{k-1}|^{p^2}}{s_{k-1}|e_{k-1}|^p |e_{k-1}|} = s_k s_{k-1}^{p-1} |e_{k-1}|^{p^2-p-1}.$$

Since $p^2 - p - 1 = 0$, we have $|e_{k-1}|^{p^2-p-1} = 1$ and

$$|r_k| = s_k s_{k-1}^{p-1}.$$

Let $\rho_k = \log|r_k|$ and $\sigma_k = \log s_k$. Then our problem is to show that the sequence defined by

$$\sigma_k = \rho_k - (p-1)\sigma_{k-1}$$

has a limit.

Let $\rho_* = \lim_{k\to\infty} \rho_k$. Then the limit σ_*, if it exists, must satisfy

$$\sigma_* = \rho_* - (p-1)\sigma_*.$$

Thus we must show that the sequence of errors defined by

$$(\sigma_k - \sigma_*) = (\rho_k - \rho_*) - (p-1)(\sigma_k - \sigma_*)$$

converges to zero. To do this we use the following easily established result from the theory of difference equations.

> *If the roots of the equation*
>
> $$x^m - a_1 x^{m-1} - \cdots - a_m = 0$$
>
> *all lie in the unit circle and* $\lim_{k\to\infty} \eta_k = 0$*, then the sequence* $\{\epsilon_k\}$ *generated by the recursion*
>
> $$\epsilon_k = \eta_k + a_1\epsilon_{k-1} + \cdots + a_m\epsilon_{k-m}$$
>
> *converges to zero, whatever the starting values* $\epsilon_0, \ldots, \epsilon_{n-1}$.

In our application $m = 1$, $\epsilon_k = (\sigma_k - \sigma_*)$, and $\eta_k = (\rho_k - \rho_*)$. The equation whose roots are to lie in the unit circle is $x + (p-1) = 0$. Since $p - 1 \cong 0.618$, the conditions of the above result are satisfied, and $\sigma_k \to \sigma_*$. It follows that the numbers s_k have a nonzero limit. In other words, two-point convergence is superlinear convergence of order p=1.618....

The generalization of this result to multipoint methods is for the most part routine. The assumption that the iteration function is identically equal to x_* whenever one of its arguments is equal to x_* implies that the Taylor series begins with the term containing $(u_1 - x_*)(u_2 - x_*)\cdots(u_n - x_*)$. The other terms of the same order that are introduced by the mean value theorem can be reduced to ones containing the product $(u_1 - x_*)(u_2 - x_*)\cdots(u_n - x_*)$ by additional Taylor expansions. The result is a recursion of the form

$$e_{k+1} = e_k e_{k-1}\cdots e_{k-n+1} r(e_k, e_{k-1}, \ldots, e_{k-n+1}).$$

where $r(0, 0, \ldots, 0)$ is finite. We can use this recursion as above to establish the convergence of the method.

The only nontrivial difference lies in establishing the rate of convergence. Proceeding as above, we obtain the following difference equation relating the $\epsilon_k = \sigma_k - \sigma_*$ and $\eta_k = \rho_k - \rho_*$:

$$\sigma_k = \eta_k - (p-1)\sigma_{k-1} - (p^2 - p - 1)\sigma_{k-2} - \cdots - (p^{n-1} - p^{n-2} - \cdots - 1)\sigma_{k-n}.$$

Here p is the positive root of the equation (4) and is easily seen to be bounded by two. If we set

$$b_i = p^i - p^{i-1} - \cdots - 1,$$

then by the result quoted above on difference equations the pth order convergence of the iteration will be established if we can show that the roots of the equation

$$h(x) = x^{n-1} + b_1 x^{n-2} + \ldots + b_{n-1} = 0 \tag{12}$$

lie in the interior of the unit circle.

Now the coefficients b_i are positive. Moreover, $b_i - b_{i+1} = p^i(2-p) > 0$. Thus the b_i satisfy

$$1 \equiv b_0 > b_1 > b_2 > \cdots > b_{n-1} > 0. \tag{13}$$

By the Eneström–Kakeya theorem (see [1]), the absolute values of the zeros of h are bounded by the largest of the ratios b_{i+1}/b_i, which are all less than one.

The following is a table of the the values of p and the magnitude q of the largest root of (12) for $n = 2, \ldots, 10$.

n	p	q
2	1.6180	0.6180
3	1.8393	0.7374
4	1.9276	0.8183
5	1.9659	0.8710
6	1.9836	0.9062
7	1.9920	0.9303
8	1.9960	0.9472
9	1.9980	0.9593
10	1.9990	0.9682

References

[1] N. Anderson, E. B. Saff, and R. S. Varga. On the Eneström–Kakeya theorem and its sharpness. *Linear Algebra and Its Applications*, 28:5–16, 1979.

[2] R. P. Brent. *Algorithms for Minimization without Derivatives.* Prentice–Hall, Englewood Cliffs, New Jersey, 1973.

[3] D. E. Muller. A method for solving algebraic equations using an automatic computer. *Mathematical Tables and Other Aids to Computation*, 10:208–215, 1956.