NUMERICAL SOLUTION OF NONLINEAR EQUATIONS IN THE ABSENCE OF THE DERIVATIVE

SUNETHRA WEERAKOON

Department of Mathematics, University of Sri Jayewardenepura, Gangodawila, Nugegoda.

Abstract: A one point iteration scheme is introduced to approximate single roots of nonlinear equations. Proposed scheme replaces the derivative of the function in Newton's Method by appropriately chosen forward or backward difference formulae. It is shown that the order of convergence of the new method is at least two and this theory is supported by computational results.

Key words: Convergence, finite difference schemes, Newton's method, nonlinear equations, root finding.

1. INTRODUCTION

Newton's method which approximates the root of a nonlinear equation using the value of the function and its derivative, in an iterative fashion, is probably the best known and most widely used algorithm. However, in many practical applications, the function is not given in closed form, rather it is the output from some computational or experimental procedure which is available. Even if the function is given by a formula, it may be too cumbersome to calculate its derivative and further Newton's method could be very expensive if the computation of the derivative requires substantially more operations than that of computing the function.

In this paper, we discuss the possibility of approximating the derivative by suitable difference approximations. It is shown that the suggested method converges quadratically and this theory is supported by computational results. It is observed that, for certain functions, suggested method can produce even better accuracy than that of Newton's method.

A similar method using forward differences for the derivative which is termed quasi Newton's method has been discussed by Dennis & Schnabel.[5] We have tested our method, which is termed Combined Finite Difference Newton's Method, for the time being, against both Newton's and quasi Newton's methods and found that for many functions our method not only converges faster, but also produces better approximations.

2. PRELIMINARY RESULTS

Definition (2.1) A function f is Lipschitz continuous with constant γ in a set A, written $f \in Lip_\gamma(A)$,

if for every $x, y \in A$, $\qquad |f(x) - f(y)| < \gamma |x - y|$ $\qquad$ ...(2.1)

Lemma (2.1): For an open interval D, let $f : D \to \mathcal{R}$ and let

$f' \in Lip_\gamma(D)$. Then for any $x, y \in D$,

$$|f(y) - f(x) - f'(x)(y-x)| \le \gamma (y - x)^2 \qquad ...(2.2)$$

Proof $\quad |f(y) - f(x) - f'(x)(y-x)|$

$\qquad = |f'(\xi)(y - x) - f'(x)(y - x)|$ $\qquad$ (by Mean Value Theorem)

$\qquad = |y - x| \, |f'(\xi) - f'(x)|$

$\qquad \le |y - x| \, \gamma |\xi - x|$ $\qquad\qquad$ (by Lipschitz continuity)

$\qquad \le \gamma |y - x|^2$ $\qquad\qquad\qquad$ (since $x \le \xi \le y$) $\qquad\qquad$ □

Lemma (2.2)[5] : For an open interval D, let $f : D \to \mathcal{R}$ and let

$f' \in Lip_\gamma(D)$. Then for any $x, y \in D$,

$$|f(y) - f(x) - f'(x)(y-x)| \le \frac{\gamma}{2} (y - x)^2 \qquad ...(2.3)$$

Proof $\quad f(y) - f(x) - f'(x)(y-x) = \int_x^y [f'(z) - f'(x)] \, dz$

$\qquad$ Let $z = x + (y-x)t$, then $dz = (y-x)dt$

Substituting these in the integral gives:

$$f(y) - f(x) - f'(x)(y - x) = \int_0^1 [f'(x + (y - x)t - f'(x)] (y - x)dt$$

Lipschitz continuity of f' together with triangle inequality gives:

$$|f(y) - f(x) - f'(x)(y-x)| \le |y - x| \int_0^1 \gamma |t(y - x)| \, dt = \gamma |y - x|^2 / 2 \qquad □$$

3. NUMERICAL SCHEMES

(3.1) Newton's Method :   Newton's algorithm to approximate the root α of the nonlinear equation $f(x) = 0$ is to start with an initial approximation x_0 sufficiently close to α and to use the one point iteration scheme

$$x_{n+1} = x_n - f(x_n)/f'(x_n) \ , \ n = 0, 1, 2, \ldots \qquad \ldots(N)$$
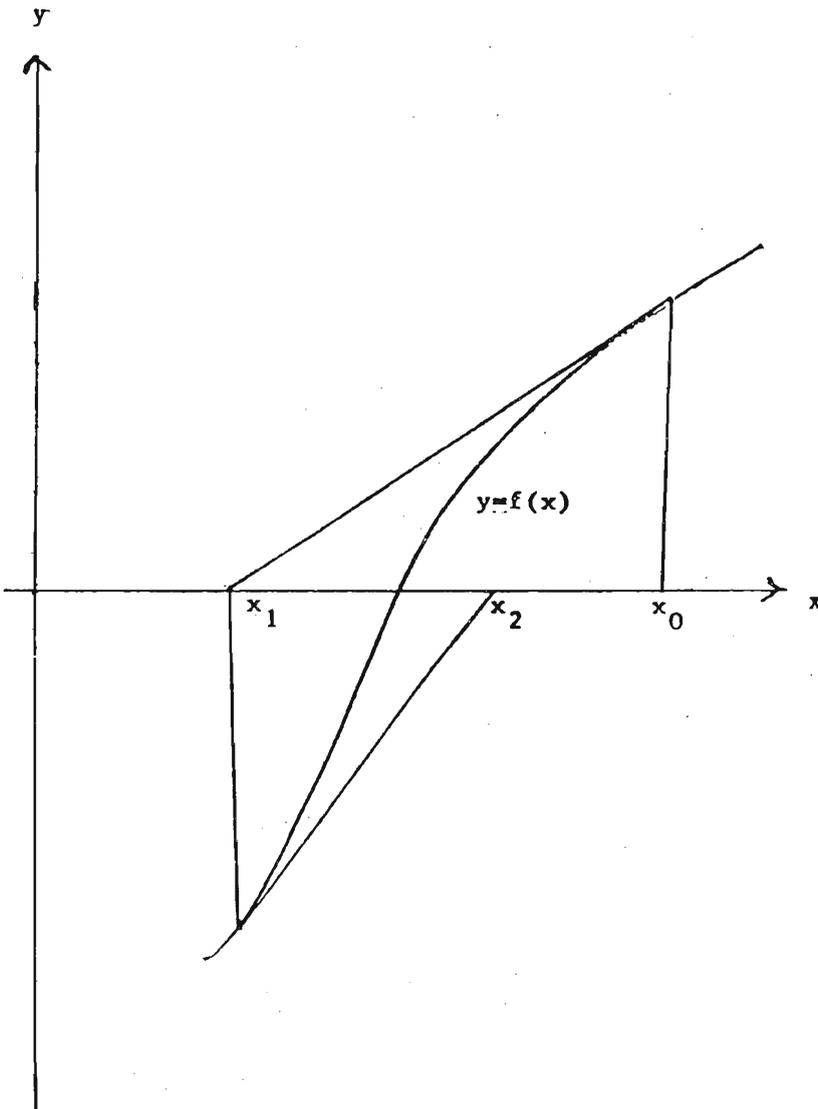

Figure 1: Newton's iterative step.

Theorem (3.1)[5]   Let $f : D \rightarrow \mathcal{R}$ for an open interval D and let
$f' \in \text{Lip}_\gamma (D)$. Assume that for some $\rho > 0$, $|f'(x)| \geq \rho$ for every $x \in D$.
If $f(x) = 0$ has a solution $\alpha \in D$, then there is some $\eta > 0$ such that :

If $|x_0 - \alpha| < \eta$, then the sequence $\{x_n\}$ generated by

$$x_{n+1} = x_n - f(x_n)/f'(x_n) \qquad n = 0 , 1 , 2 , ...$$

exists and converges to α . Furthermore, for $n = 0 , 1 , ...$

$$|x_{n+1} - \alpha| \leq \frac{\gamma}{2\rho} |x_n - \alpha|^2 \tag{3.1}$$

(3.2) Combined Finite Difference Newton's Method : The combined Finite
Difference Newton's method is the new method suggested in this paper. When x_0
is sufficiently close to the root α, it can be given by the one point iteration scheme

$$x_{n+1} = x_n - f(x_n)/b_n , \quad n = 0, 1, ...$$

$$b_n = \begin{cases} \dfrac{f(x_n + h_n) - f(x_n)}{h_n} , \text{ if } |f(x_n) + h_n)| < |f(x_n)| \\[4mm] \dfrac{f(x_n) - f(x_n - h_n)}{h_n} , \text{ if } |f(x_n - h_n)| < |f(x_n)| \end{cases} \quad ...\textbf{(FDN)}$$
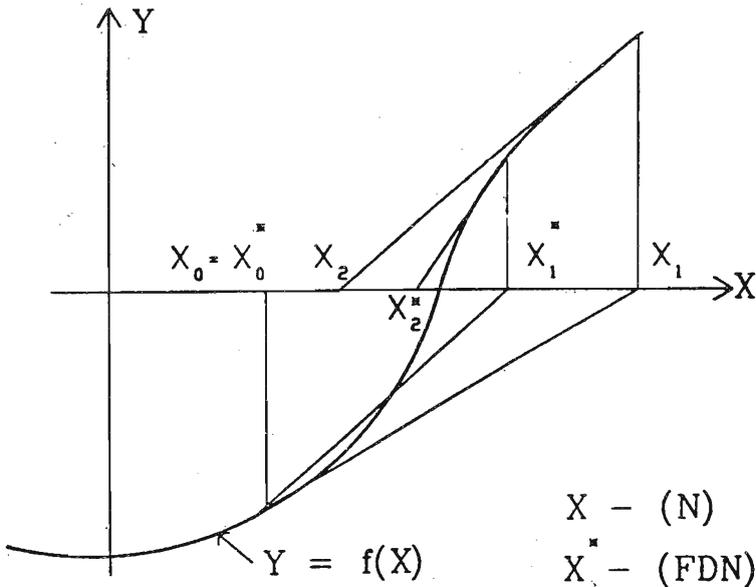

Figure 2: Convergence of (FDN) and (N) Methods.

4. ANALYSIS OF CONVERGENCE

Theorem (4.1)   Let $f : D \to \mathcal{R}$ for an open interval D and let $f' \in \mathrm{Lip}_\gamma(D)$. Assume that for some $\rho > 0$, $|f'(x)| \geq \rho$ for every $x \in D$. If $f(x) = 0$ has a simple root $\alpha \in D$, then the combined finite difference Newton's method defined by (FDN) is quadratically convergent to the root α for x_0 sufficiently close to α and h_n sufficiently small.

Lemma (4.1)   Let $f : D \to \mathcal{R}$ for an open interval D and let $f' \in \mathrm{Lip}_\gamma(D)$. Assume that for some $\rho > 0$, $|f'(x)| \geq \rho$ for every $x \in D$. Then b_n defined by (FDN) satisfies the inequality:

$$|b_n^{-1}| \leq \rho^{-1}$$

Proof $h_n |b_n| = |f(x \pm h_n) - f(x_n)| = |f'(\xi)| \, h_n$ for some $\xi \in D$

(by Mean Value Theorem)

i.e. $|b_n| = f'(\xi)| \geq \rho$ (since $\xi \in D$)

Hence $|b_n^{-1}| \leq \rho^{-1}$ □

Lemma (4.2)   Let $f : D \to \mathcal{R}$ for an open interval D and let $f' \in \mathrm{Lip}_\gamma(D)$. Then b_n defined by (FDN) satisfies the following inequality:

$$|b_n - f'(x_n)| \leq \frac{\gamma h_n}{2} \qquad \text{for every } x_n \pm h_n \in D$$

Proof $|b_n - f'(x_n)| = |f(x_n + h_n) - f(x_n) - f'(x_n)h_n| \, |h_n^{-1}|$

$= |f(x_n + h_n) - f(x_n) - f'(x_n)(x_n + h_n - x_n)| \, |h_n^{-1}|$

$\leq \gamma (|x_n + h_n - x_n|^2/2)h_n^{-1}$ (by Lemma (2.2))

$= \gamma h_n / 2$ □

Proof of Lemma (4.2) for $b_n = \dfrac{[f(x_n) - f(x_n - h_n)]}{h_n}$ is similar.

Proof of Theorem (4.1):

$$\text{Let} \quad e_n = |\alpha - x_n|$$

We have $\qquad x_{n+1} = x_n - f(x_n)/b_n$ and $f(\alpha) = 0$

Thus $\qquad x_{n+1} - \alpha = x_n - \alpha - f(x_n)/b_n$

$$= b_n^{-1}\{f(\alpha) - f(x_n) - b_n(\alpha - x_n)\}$$

$$= b_n^{-1}\{f(\alpha) - f(x_n) - f'(x_n)(\alpha - x_n) + (f'(x_n) - b_n)(\alpha - x_n)\}$$

$$\Rightarrow \quad e_{n+1}$$

$$\le |b_n^{-1}| \{ |f(\alpha) - f(x_n) - f'(x_n)(\alpha - x_n)| + |(f'(x_n) - b_n)|e_n\}$$

$$\le \rho^{-1}\{\frac{\gamma}{2}e_n^2 + |(f'(x_n) - b_n)|e_n\} \quad \text{(by Lemmas (4.1) \& (2.2))}$$

$$\le \rho^{-1}\{\frac{\gamma}{2}e_n^2 + \gamma h_n e_n/2\} \qquad \text{(by Lemma (4..2))}$$

$$\le (\gamma/\rho)e_n^2 \qquad (\text{if } h_n < e_n) \qquad\qquad\qquad ...(4.1) \qquad \square$$

Note : It is possible to establish the second order convergence of (FDN) using only Lemma (2.1) instead of Lemma (2.2). However, the constant in (4.1) then becomes $2(\gamma/\rho)$ instead of γ/ρ.

5. COMPUTATIONAL RESULTS AND DISCUSSION

We applied both methods (N) and (FDN) to several functions using different initial values and stepsizes. Tables 1 and 2 contain summaries of the results obtained.

(5.1) Analysis of Computational Results : Every function tested with various initial values and stepsizes not only supports the proven theory of quadratic convergence of (FDN) but goes beyond that and suggests the order of convergence of the proposed method could be even higher! Also note that in all cases, error in (FDN) is smaller than that of (N). Though this has nothing to do with the order of convergence, we can visualize this geometrically. In other words, while we can reach arbitrarily close to the root by adjusting the stepsize in (FDN) , (N) doesn't provide such flexibility. This is evident in both tables 1 and 2.

Table 1 : Comparison of (N) and (FDN) after 5 iterations.

Function f(x)	Step Size h		Initial Guess x_0	Absolute Error e_5		Average Order of Convergence	
				N	FDN	N	FDN
$(x-1)^3-1$	10^{-1}	$\lvert f(x)\rvert$	1.5	2.455×10^{-6}	1.05×10^{-10}	1.83	1.88
	10^{-2}	$\lvert f(x)\rvert$		same	1.06×10^{-6}	1.832	1.839
	10^{-12}	$\lvert f(x)\rvert$		same	2.3×10^{-7}	1.83	2.07
	10^{-11}			same	2.454×10^{-6}	1.83	1.89
x^3-10	10^{-3}	$\lvert f(x)\rvert$	1.5	3.5×10^{-14}	2.5×10^{-14}	1.96	1.96
	10^{-2}	$\lvert f(x)\rvert$		same	0.0	1.96	1.99
	10^{-13}		0.4	2.13	2.13	1.89	1.97
$x^2\sin^2x+e^{(x2\cos x+\sin x)}-28$	10^{-5}		4.0	2.457×10^{-2}	1.0316×10^{-3}	1.154	1.8809
	10^{-8}		4.0	same	1.0344×10^{-3}	same	1.8801
$xe^{x2}-\sin^2x+3\cos x+5$	10^{-8}		-1.5	1.881×10^{-5}	0.0	1.003	1.9524
			-2.0	3.119×10^{-3}	1.8×10^{-3}	0.847	0.982
$16x+(1+x)^{-20}-1$	10^{-5}		0.01066	1.85×10^{-4}	0.0	1.257	1.898

Table 2 : Comparison of (N) and (FDN) after 10 iterations.

Function f(x)	Step Size h	Initial Guess x_0	Absolute Error e_{10}		Average Order of Convergence	
			N	FDN	N	FDN
x^3-10	10^{-13}	0.4	2.24×10^{-8}	7.63×10^{-9}	1.89	1.98
Arctanx	10^{-9}	1.39	5.7×10^{-13}	1.2×10^{-14}	2.76	2.80
$x^2 \sin^2 x + e^{(x^2 \cos x + \sin x)} - 28$	10^{-5}	4.0	4.26×10^{-2}	0.0	1.52	>1.98
$xe^{x^2} - \sin^2 x + 3\cos x + 5$	10^{-8}	-1.5	1.0×10^{-8}	0.0	1.29	>1.95
		-2.0	2.32×10^{-6}	0.0	1.001	>1.98

(5.2) Number of Function Evaluations : One might be tempted to argue that (FDN) as presented require one more function evaluation than that of (N). When we know the behaviour of the function, which is the case most of the time, we can avoid evaluating both $f(x_n + h_n)$ and $f(x_n - h_n)$ at each step by indicating which formula to be used. For example, if $f(x)$ is an increasing function, we know that the backward difference formula should be used when $f(x) > 0$ and it is the forward difference formula when $f(x) < 0$. If we incorporate this knowledge in programming, only one fresh function evaluation is necessary for each iteration.

(5.3) Choice of Stepsize : We established the second order convergence of the method imposing the condition $h_n < e_n$ on the stepsize. Clearly it is not a necessary condition but it is sufficient. However, we have to make sure that this requirement is met computationally, to ensure second order convergence.
Note that

$$| f(x_n) - f(\alpha) | \; < \beta \; | x_n - \alpha | \qquad (\text{if, } f \in \text{Lip}_\beta(D))$$

$$\text{i.e. } \beta^{-1} \, | f(x_n) | \; < \; e_n$$

Thus if we choose $h_n \leq \beta^{-1} \, | f(x_n) | \quad (< e_n)$, second order convergence is ensured.

In some cases, we can reach closer to the root by adjusting the stepsize. This does not necessarily mean that we have to reduce the stepsize. Observe that in cases of $f(x) = x^3 - 10$ and $f(x) = (x-1)^3 - 1$, convergence was faster for larger stepsize. We can suggest that it is advisable to select the largest possible stepsize, bearing in mind that conditions $x_n \pm h_n \in D$ and $h_n < e_n$ are also required for second order convergence.

We have proved in theorem (4.1) that (FDN) is at least second order convergent. In fact, the computational results overwhelmingly suggest that (FDN) converges faster than Newton's method in all cases. Hence it is best to use (FDN) for root finding not only in the absence of the derivative but wherever Newton's method is applicable.

Now that we have established the Finite Difference Newton's method by providing computational evidence and sound theoretical support, the obvious next directions would be to apply the method to functions of several variables and to unknown functions whose values at required points in the domain could be generated by some experimental or computational procedure.

References

1. Atkinson Kendall E. (1988). *An introduction to numerical analysis,* John Wiley, N.Y.

2. Atkinson Kendall E. (1985). *Elementary numerical analysis*, John Wiley, N.Y.

3. Amos Ehrlich. (1991). The Lipschitz condition without tears. *International Journal for Mathematical Education in Science and Technology* **22**(4): 519-522.

4. Burden Richard L & Faires Douglas J. (1993). *Numerical analysis* - PWS-Kent Publishing Co, Boston.

5. Dennis J. E. & Schnable Robert B. (1983). *Numerical methods for unconstrained optimization and nonlinear equations.* Prentice Hall, Inc.

6. Isaacson E. & Keller H. B. (1996). *Analysis of numerical methods*, John Wiley, N.Y.

7. Jain M.K. (1971). *Numerical analysis for scientists and engineers*, S. B. W. Publishers, Delhi.

8. Ortega J. & Rheinboldt W. (1970). *Iterative solutions of nonlinear equations in several variables*, Academic Press, New York.

9. Rheinboldt Werner C. (1987). *Methods of solving systems of linear equations,* SIAM.

10. Wilkinson J. (1963). *Rounding errors in algebraic processes*, Prentice Hall, Englewood Cliffs, N.J.