



A slight generalization of Steffensen Method for Solving Non Linear Equations

  10.56238/tfisdwv1-142

Eder Marinho Martins

Department of Mathematics, Federal University of Ouro Preto, MG, Brazil
Email: eder@ufop.edu.br

Geraldo Cesar Gonçalves Ferreira

Department of Mathematics, Federal University of Ouro Preto, MG, Brazil
Email: geraldocesar@ufop.edu.br

Thais Ester Gonçalves

Department of Mathematics, Federal University of Ouro Preto, MG, Brazil
Email: thais.ester@aluno.ufop.edu.br

ABSTRACT

If In this article, we present an iterative method to find simple roots of nonlinear equations, that is, to solving

an equation of the form $f(x) = 0$. Different from Newton's method, the method we propose do not require evaluation of derivatives. The method is based on the classical Steffensen's method and it is a slight modification of it. The proofs of theoretical results are stated using Landau's Little o notation and simple concepts of Real Analysis. We prove that the method converges and its rate of convergence is quadratic. The method present some advantages when compared with Newton's and Steffensen's methods as illustrated by numerical tests given.

Mathematics subject classification: 65-02, 65-11, 65B99

Keywords: Iteration Method; Quadratic Convergence; Nonlinear Equations; Steffensen Method.

1 INTRODUCTION

Iterative methods for solving nonlinear real equations of the form $f(x) = 0$ have been widely studied by many researchers around the world. Newton's (or Newton-Raphson's) method certainly is the best known iterative method and studied in any numerical calculus course. From a initial guess x_0 , it is defined the sequence (x_n) given by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad (1.1)$$

when $f'(x_n) \neq 0$

(it means the sequence is well defined). It is shown (see, for example, [3]), under certain hypothesis, if x_0 is chosen close enough to p , where $f(p) = 0$, then $\lim x_n = p$. On the one hand, Newton's method is very efficient, because its convergence is quadratic, but on the other hand it is necessary to compute derivatives, what can be computationally expensive. An alternative idea is to approximate the derivative in the method in some way. Secant's (see [1]), Steffensen's (see [2]) and Kurchatov's (see [5]) method applies this idea (free derivative methods) and probably are the best known iterative algorithms free

of derivatives. As is known in the literature, Secant's method has rate of convergence given by the golden ratio number, Steffensen's and Kurchatov's have quadratic convergence. Based on these ideas, many reserachers have been worked to obtain free derivative methods for solving non linear equations (see, for example, [4, 6–9]).

In this paper, we present an iterative method for finding simple root of a non linear equation based on Steffesen Method. In spite of the simplicity of our ideas, we have not been able to find any reference in the literature as we do here. The algorithm purposed is given by

$$x_{n+1} = x_n - \frac{g \circ f(x_n) \cdot f(x_n)}{f(x_n + g \circ f(x_n)) - f(x_n)} \quad (1.2)$$

where g is a continuous differentiable function that has in the origin an isolated zero. Observe that if g is the identity function, one has the classical Steffensen method. In this paper, we refer to (1.2) as g -Steffensen method.

Our main strategy is to use Landau's Little o notation and its algebra which simplifies the way of writing proofs. This paper is organized as follows. In Section 2, we present the little o notation. For completeness, in Proposition 2.1 we enumerate simple, but important properties that permit us describe Little o notation's algebra. In Section 3, we proof two theorems that ensure convergence and rate of convergence. In Section 4, are given some numerical tests and comments. We present some examples when (1.2) is effective for some examples of g functions compared with Steffensen's and Newton's method.

2 A NOTE ON LITTLE O NOTATION

For completeness, we defines what is the little o notation and the algebra utilized.

Definition 2.1. Let X be a nonempty subset of \mathbb{R} and a a limit point of X . For given functions

$f, g : X \rightarrow \mathbb{R}$, we say that $f(x) = o(g(x))$, or just $f = o(g)$, as $x \rightarrow a$, if $f \in o(g)$ where $o(g) = \{h : X \rightarrow \mathbb{R} : \forall \varepsilon \exists \delta > 0 \text{ such that } |h(x)| < \varepsilon |g(x)| \text{ for all } x \in (a - \delta, a + \delta) \setminus \{a\}\}$.

In a similar way, one can define

Definition 2.2. Let $X \subset \mathbb{R}$ be unbounded above. For given functions $f, g : X \rightarrow \mathbb{R}$, we say that $f(x) = o(g(x))$, or just $f = o(g)$, as $x \rightarrow \infty$, if $f \in o(g)$, where

$o(g) = \{h : X \rightarrow \mathbb{R} : \forall \varepsilon \exists M > 0 \text{ such that } |h(x)| < \varepsilon |g(x)| \text{ for all } x > M \}$.

The proof of proposition below is straightforward and it is omitted.

Proposition 2.1. Let f, g, h, F and G be real functions. Also for $x \rightarrow a$ and $x \rightarrow \infty$, if

(i) $c \neq 0$ and $f = o(F)$, then $cf = o(F)$;

- (ii) $f = o(F)$ and $g = o(G)$, then $f \cdot g = o(FG)$;
- (iii) $f = o(g)$ and $g = o(h)$, then $f = o(h)$;
- (iv) $f = o(F)$ and $g = o(G)$, then $f + g = o(H)$ as $x \rightarrow a$, where $H = \max\{F, G\}$.

According to Proposition 2.1 it is possible to define an algebra for Little o notation.

Definition 2.3. Let f, g be real functions. One can define the operations (also for $x \rightarrow a$ or $x \rightarrow \infty$):

- (i) $c \cdot o(f) = o(f)$, if $c \neq 0$;
- (ii) $o(f) \cdot o(g) = o(fg)$;
- (iii) $f \cdot o(g) = o(fg)$;
- (iv) $o(f) + o(g) = o(h)$, where $h = \max\{f, g\}$.

3 CONVERGENCE OF G-STEFFENSEN METHOD

Theorem 3.1. Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$ be continuously differentiable real functions. Assume that p is a isolated zero of f such that $f'(p) \neq 0$. Suppose that g has an isolated zero in origin. If f' is continuous, then there is a neighborhood V of p such that the sequence (x_n) produced by (1.2), where $x_0 \in V$, converges to p .

Proof. Let φ be the real function given by

$$\varphi(x) = x - \frac{f(x)}{g(x)}$$

where

$$\rho(x) = \begin{cases} \frac{f(x+g(f(x))) - f(x)}{g(f(x))}, & \text{if } x \neq p \\ f'(p), & \text{if } x = p \end{cases}$$

It is straightforward to see that ρ is a continuous function and that fixed points of ρ are roots of f . From Taylor Series expansion and Definition 2.3, as $x \rightarrow p$, we have:

$$\begin{aligned} \rho'(x) &= f''(x) + o(1) + \frac{f'(x)}{g \circ f(x)} + f''(x) + o(1) \cdot (g \circ f)'(x) \\ &- (g \circ f)'(x) \frac{f(x)}{g \circ f(x)} + \frac{f'(x)}{2} + o(1) \\ &= f''(x) + \frac{1}{2} f''(x)(g \circ f)'(x) + ((g \circ f)'(x)) \cdot o(1). \end{aligned}$$

Therefore

$$\lim_{x \rightarrow p} \rho'(x) = f''(p) + \frac{1}{2} f''(p) \cdot (g \circ f)'(p).$$

On other hand, we can write

$$\begin{aligned} \frac{\rho(x) - \rho(p)}{x - p} &= \frac{f(x) + \frac{f''(x)}{2} g \circ f(x) + o(g \circ f(x)) - f(p)}{x - p} \\ &= \frac{f(x) - f(p)}{x - p} + \frac{1}{2} f''(x) \cdot \frac{g \circ f(x) - g \circ f(p)}{x - p} + \frac{g \circ f(x) - g \circ f(p)}{x - p} o(1). \end{aligned}$$

Then

$$\rho(p) = f(p) + \frac{1}{2} f''(p) \cdot (g \circ f)'(p).$$

According (3.1), ρ is a C1 function, as well φ . This way, we have $\varphi'(p) = 0$. This ensure the existence of V and the theorem is proved.

The quadratic converge can be obtained as consequence of Theorem 1 of [6], but we present another prove for completeness.

Theorem 3.2. The rate of convergence of Algorithm given in (1.2) is quadratic.

Proof. Let p be such that $f(p) = 0$ and $e_n = x_n - p$. From (1.2), we have

$$\begin{aligned} e_{n+1} &= e_n - \frac{g \circ f(x_n) \cdot f(x_n)}{f(x_n + g \circ f(x_n)) - f(x_n)} \\ &= e_n - \frac{g \circ f(x_n) \cdot f(x_n)}{f(x_n) + f(x_n) \cdot g \circ f(x_n) + \frac{f''(x_n)}{2} \cdot g \circ f(x_n) + o(g \circ f(x_n)) - f(x_n)} \\ &= e_n - \frac{f(x_n)}{f(x_n) + \frac{f''(x_n)}{2} \cdot g \circ f(x_n) + o(g \circ f(x_n))}. \end{aligned}$$

Since

$$f(x_n) = f(p) + f'(p)e_n + \frac{f''(p)}{2}e_n^2 + o(e_n^2) = f(p)e_n + \frac{f''(p)}{2}e_n^2 + o(e_n^2)$$

and $f(x_n) = f(p) + f'(p)e_n + o(e_n)$, one has

$$\begin{aligned} e_{n+1} &= e_n - \frac{f(p)e_n + \frac{f''(p)}{2}e_n^2 + o(e_n^2)}{f(p) + f'(p)e_n + o(e_n) + g \circ f(x_n) \frac{f''(x_n)}{2} + o(1)} \\ &= \frac{\frac{f''(p)}{2}e_n^2 + o(e_n^2) + g \circ f(x_n) \frac{f''(x_n)}{2} + o(1)}{f(p) + f'(p)e_n + o(e_n) + g \circ f(x_n) \frac{f''(x_n)}{2} + o(1)} e_n \end{aligned}$$

On other hand,

$$\alpha \circ \xi(x^u) = \alpha \circ \xi(b) + (\alpha \circ \xi)'(b)\epsilon^u + o(\epsilon^u) = (\alpha \circ \xi)'(b)\epsilon^u + o(\epsilon^u)$$

then one can write

$$e_{n+1} = \frac{\frac{f''(p)}{2}e_n^2 + o(e_n^2) + (g \circ f)'(p)e_n + o(e_n) \cdot \frac{f''(x_n)}{2} + o(1)}{f(p) + f'(p)e_n + o(e_n) + ((g \circ f)'(p)e_n + o(e_n)) \cdot \frac{f''(x_n)}{2} + o(1)}$$

that gives

$$e_{n+1} = \frac{\frac{f''(p)}{2} + (g \circ f)'(p) + o(1)}{f(p) + f'(p)e_n + o(e_n) + ((g \circ f)'(p)e_n + o(e_n)) \cdot \frac{f''(x_n)}{2} + o(1)}$$

It follows that

$$\lim \frac{e_{n+1}}{e_n^2} = \frac{1}{f'(p)} \cdot \frac{f''(p)}{2} + (g \circ f)'(p) \cdot \frac{f''(x_n)}{2} = \frac{1}{f'(p)} \cdot \frac{f''(p)}{2} + g'(0)f'(p) \cdot \frac{f''(p)}{2} . 5$$

That is, the g -Steffensen iterative method has quadratic convergence.

4 NUMERICAL TESTS

In this section we present some numerical tests using iteration formulae (1.2). For these tests, we chose six g functions: $g_1(x) = \sin(x)$, $g_2(x) = e^x - 1$, $g_3(x) = x^2$, $g_4(x) = \cos(x) - 1$, $g_5(x) = \tan(x)$ and $g_6(x) = e^{-x} - 1$. In all examples, we looked for the root p in the interval $[a, b]$ and chose the initial shoot x_0 belonging to $[a, b]$. In the tables presented, n indicates the number of iterations and x_n illustrates an approximation of p . Comparisons with Steffensen's and Newton's method are also commented.

Examples of functions that do not converge for Steffensen method

Examples below are not convergent when we use Steffensen method, that is, when g is the identity function. The g -Steffesen method is convergent for many g 's.

Example 4.1. $f(x) = \sin^2(x) - x^2 + 1$, $[a, b] = [0, 3]$, $x_0 = 3$.

The g -Steffensen method is convergent for all g 's we chose (see Table 4.1).

Table 4.1: $f(x) = \sin^2(x) - x^2 + 1$

	n	x_n	$ f(x_n) $
g_1	7	1.4044916482153411	$3.3306690738754696 \times 10^{-16}$
g_2	7	1.4044916482153411	$3.3306690738754696 \times 10^{-16}$
g_3	12	1.4044916482773504	$1.5393597507795675 \times 10^{-10}$
g_4	4	1.4044916488265187	$1.5172312295419488 \times 10^{-9}$
g_5	11	1.4044916482153413	$4.440892098500626 \times 10^{-16}$
g_6	80	1.4044916482153411	$3.3306690738754696 \times 10^{-16}$

Example 4.2. $f(x) = x^3 - x - 1$, $[a, b] = [0, 2]$, $x_0 = 1$.

Among all g functions we chose, just for g_2 the algorithm (1.2) is not convergent (see Table 4.2).

Table 4.2: $f(x) = x^3 - x - 1$

	n	x_n	$ f(x_n) $
g_1	11	1.324717957244746	$2.220446049250313 \times 10^{-16}$
g_3	5	1.3247179573200405	$3.211033661187912 \times 10^{-10}$
g_4	5	1.3247179573118653	$2.862388104318825 \times 10^{-10}$
g_5	28	1.324717957244746	

g_6	8	1.324717957244746	$2.220446049250313 \times 10^{-16}$
			$2.220446049250313 \times 10^{-16}$

Example 4.3. $f(x) = e^{1-x} - 1$, $[a, b] = [0, 3]$, $x_0 = 3$.

The method is convergent for g_1, g_4, g_5, g_6 (see Table 4.3) and it is divergent for g_2 and g_3 .

Table 4.3: $f(x) = e^{1-x} - 1$

	n	x_n	$ f(x_n) $
g_1	5	1.0	0.0
g_4	11	0.9999999999188026	$8.119727112898545 \times 10^{-11}$
g_5	6	1.0	0.0
g_6	24	0.9999999999999999	0.0

EXAMPLES OF FUNCTIONS THAT DO NOT CONVERGE FOR NEWTON'S AND STEFFENSEN'S METHOD

We present some examples when both Newton's and Steffensen's method do not converge, but the g -Steffensen method is convergent for some choices of g .

Example 4.4. $f(x) = x^3 - 2x + 2$, $[a, b] = [-3, 1]$, $x_0 = 1$.

In this example, g -Steffensen method is convergent for g_1, g_3 and g_4 (see Table 4.4) and it is divergent for g_2, g_5 and g_6 .

Table 4.4: $f(x) = x^3 - 2x + 2$

	n	x_n	$ f(x_n) $
g_1	12	-1.7692923542386314	0.0
g_3	28	-1.7692923543026569	$4.73223682462276 \times 10^{-10}$
g_4	27	-1.76929235421728	$1.5781242979073795 \times 10^{-10}$

Example 4.5. $f(x) = \arctg(x - 2)$, $[a, b] = [0, 3.5]$, $x_0 = 3.5$.

The g -Steffensen method is convergent for g_4 and g_6 (see Table 4.5) and it is divergent for g_1, g_2, g_3 and g_5 .

Table 4.5: $f(x) = \text{arctg}(x - 2)$

	n	x_n	$ f(x_n) $
g_4	6	2.0000000000000001	$8.881784197001252 \times 10^{-16}$
g_6	4	2	0

Example 4.6. $f(x) = x^5 - x + 1$, $[a, b] = [0, 3]$, $x_0 = 3$.

If we use g_1 , g_4 and g_6 , g -Steffensen method is convergent (see Table 4.6), but it is divergent for g_2 , g_3 and g_5 .

Table 4.6: $f(x) = x^5 - x + 1$

	n	x_n	$ f(x_n) $
g_1	30	-1.1673039782614187	$6.661338147750939 \times 10^{-16}$
g_4	8	-1.1673039788241997	$4.6617254501057914 \times 10^{-9}$
g_6	22	-1.1673039782614187	$6.661338147750939 \times 10^{-16}$

OTHER CASES

Example 4.7. $f(x) = 0.5x^3 - 6x^2 + 21.5x - 22$, $[a, b] = [0, 3]$, $x_0 = 3$.

Newton's method is divergent, Steffensen's method is convergent to a root not in interval $[0,3]$ and Algorithm (1.2) converges to a root in $[0, 3]$ for g_2 and g_4 (see Table 4.7); it is convergent to a root outside $[0, 3]$ for g_1 and g_3 , and it is divergent for g_5 and g_6 .

Table 4.7: $f(x) = 0.5x^3 - 6x^2 + 21.5x - 22$

	n	x_n	$ f(x_n) $
g_2	20	1.7639320225002113	$7.105427357601002 \times 10^{-15}$
g_4	5	1.7639320224170847	$4.156319732828706 \times 10^{-10}$

Example 4.8. $f(x) = \cos(x)$, $[a, b] = [0, 3.5]$, $x_0 = 3.5$.

In this example, Newton's method and Steffensen's method converges to a root outside of the interval considered. The method g -Steffesen method is convergent just for g_5 as illustrated in Table 4.8.

Table 4.8: $f(x) = \cos x$

	n	x_n	$ f(x_n) $
g_5	5	1.5707963267948966	$6.123233995736766 \times 10^{-17}$

Example 4.9. $f(x) = 10xe^{-x} - 1$, $[a, b] = [0, 3]$, $x_0 = 3$.

In this example, Newton's method and Steffensen's method are divergent. The method presented in this article is convergent, for g_5 with 8 iterations for the root belonging to the interval, $x_n = 1.67963061042845$, resulting in $|f(x_n)| = 2.220446049250313 \times 10^{-16}$ (see Table 4.9). For g_1, g_2, g_3, g_4 and g_6 , the method is divergent.

Table 4.9: $f(x) = 10xe^{-x} - 1$

	n	x_n	$ f(x_n) $
g_5	8	1.67963061042845	$2.220446049250313 \times 10^{-16}$

ACKNOWLEDGMENTS

Authors thanks Federal University of Ouro Preto and first and third authors thanks FNDE/MEC for partial support.

REFERENCES

- [1] D. Pedro, A note on the convergence of the secant method for simple and multiple roots, *Applied mathematics letters*, 16:8 (2003), 1211–1215.
- [2] J. F. Steffensen, Remarks on iteration, *Scandinavian Actuarial Journal*, 1: (1933), 64-72.
- [3] Richard L. Burden and J. Douglas Faires, *Numerical methods*, Thomson, 2003.
- [4] L. Piscoran and D. Miclaus, A New Steffensen-Homeier Iterative Method for Solving Nonlinear Equations, *Investigaci' on Operacional*, 40:1 (2019), 74-80.
- [5] S. M. Shakhno, On a Kurchatov's method of linear interpolation for solving nonlinear equations, *Proceedings in Applied Mathematics and Mechanics*, 4:1 (2004), 650-651.
- [6] V. Candela and R. Peris, A class of third order iterative Kurchatov–Steffensen (derivative free) methods for solving nonlinear equations, *Applied Mathematics and Computation*, 350 (2019), 93-104.
- [7] X. Wu and D. Fu, New high-order convergence iteration methods without employing derivatives for solving nonlinear equations, *Computers & Mathematics with Applications*, 41:3-4 (2001), 489–495.
- [8] X. Wu and H. Wu, On a class of quadratic convergence iteration formulae without derivatives, *Applied Mathematics and Computation*, 107:2-3 (2000), 77-80.
- [9] Z. Wang and X. Wu, A semi-local convergence theorem for a robust revised Newton's method, *Computers & Mathematics with Applications*, 58:7 (2009), 1320-1327.