Applied Mathematics Letters 25 (2012) 1881–1886

Contents lists available at SciVerse ScienceDirect

Applied Mathematics Letters

journal homepage: www.elsevier.com/locate/aml

A cubically convergent Steffensen-like method for solving nonlinear equations

Jovana Džunić *, Miodrag S. Petković

Faculty of Electronic Engineering, Department of Mathematics, University of Niš, 18000 Niš, Serbia

ARTICLE INFO

Article history:
Received 9 November 2011
Received in revised form 22 February 2012
Accepted 22 February 2012

Keywords:
Nonlinear equations
Iterative methods
Multipoint methods with memory
Acceleration of convergence
R-order of convergence
Computational efficiency

ABSTRACT

A derivative free method for solving nonlinear equations of Steffensen's type is presented. Using a self-correcting parameter, calculated by using Newton's interpolatory polynomial of second degree, the *R*-order of convergence is increased from 2 to 3. This acceleration of the convergence rate is attained without any additional function calculations, which provides a very high computational efficiency of the proposed method. Another advantage is the convenient fact that this method does not use derivatives. Numerical examples are included to confirm the theoretical results and high computational efficiency.

1. Introduction

The main principle in constructing iterative algorithms for solving nonlinear equations is to achieve as high as possible a convergence rate with a fixed number of function evaluations per iteration. In this work we use a self-correcting parameter to improve the quadratically convergent Steffensen-like method and its modification with memory, both proposed by Traub in [1]. This parameter is calculated in each iteration by employing available data from the current and the previous iteration and so no additional computational cost is required. In this way we construct a new efficient method with memory with the *R*-order of convergence at least 3. The improved method is free of derivatives, which is another advantage.

Let α be a simple real zero of a real function $f : D \subset \boldsymbol{R} \to \boldsymbol{R}$ and let x_0 be an initial approximation to α. In his book [1], Traub considered the iterative function

$$\Phi(x) = \Phi(x, \gamma) = x - \frac{\gamma f(x)^2}{f(x + \gamma f(x)) - f(x)} \tag{1}$$

where $\gamma \neq 0$ is a real constant. Introducing $u(x) = f(x)/f'(x)$ and expanding the denominator in (1) in a geometrical series, we arrive at the relation

$$\Phi(x) - \alpha = (1 + \gamma f'(x))c_2(x)u(x)^2 + O(u(x)^3). \tag{2}$$

In particular, choosing $x = \alpha$ it follows that $\Phi(\alpha) = \alpha$ and $\Phi'(\alpha) = 0$, which means that (1) defines at least a second-order iteration according to the Schröder–Traub theorem [1, Theorem 2.2]. Note that $\gamma = 1$ gives the well-known method of Steffensen [2]:

$$x_{k+1} = x_k - \frac{f(x_k)^2}{f(x_k + f(x_k)) - f(x_k)} \quad (k = 0, 1, \ldots).$$

* Corresponding author.
E-mail address: jovana.dzunic@elfak.ni.ac.rs (J. Džunić).

0893-9659/$ – see front matter
doi:10.1016/j.aml.2012.02.055

Taking $\gamma = -1/f'(x)$ and having in mind that $f(x) = O(x-\alpha)$, we conclude from (2) that

$$\Phi(x) - \alpha = O((x-\alpha)^3). \tag{3}$$

For this particular choice of γ, from (1) we get the Newton-secant iterative method

$$x_{k+1} = x_k - \frac{u(x_k)f(x_k)}{f(x_k) - f(x_k - u(x_k))} \quad (k = 0, 1, \ldots) \tag{4}$$

which is, according to (3), of third order. This iterative function was derived by Traub and can be regarded as a two-point method. Note that (4) requires three function evaluations. Using the Ostrowski–Traub formula for the efficiency index of an iterative method (IM):

$$E(\mathrm{IM}) = r^{1/\theta}, \tag{5}$$

where r is the order of (IM) and θ is the required number of function evaluations per iteration, we calculate $E(4) = 3^{1/3} \approx 1.442$. The third-order one-point methods with the highest computational efficiency, such as Halley's method (23) and Ostrowski's method (24), possess the same efficiency index.

Regarding the construction of the two-point method (4) where the parameter γ is replaced by $-1/f'(x)$, a reasonable question arises; can we increase the order of convergence of (1) dealing with some approximation $\bar{f}'(x)$ of $f'(x)$ and taking $\gamma = -1/\bar{f}'(x)$? This substitution must be cost-preserving, that is, γ should be calculated using only available information.

Traub [1] showed that Steffensen-like method (1) can be somewhat improved by the reuse of information from the previous iteration. Approximating f' by the secant

$$\bar{f}'(x_k) = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}, \tag{6}$$

Traub proposed the following method with memory:

$$\begin{cases} \gamma_0 \text{ is given}, \quad \gamma_k = -\dfrac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \text{ for } k \geq 1, \\ x_{k+1} = x_k - \dfrac{\gamma_k f(x_k)^2}{f(x_k + \gamma_k f(x_k)) - f(x_k)}, \end{cases} \quad (k = 0, 1, \ldots) \tag{7}$$

with the R-order of convergence at least $1+\sqrt{2} \approx 2.414$. A similar approach was applied to higher order multipoint methods in [3,4].

In this work we show that the iterative method (7) can be additionally accelerated without increasing the computational cost. The main idea in constructing a higher order method consists of the calculation of the parameter $\gamma = \gamma_k$ as the iteration proceeds using a better approximation to $f'(x_k)$ related to (6).

2. The improved method with memory

The main idea in constructing methods with memory consists of the calculation of the parameter $\gamma = \gamma_k$ as the iteration proceeds by using the formula $\gamma_k = -1/\bar{f}'(\alpha)$ for $k = 1, 2, \ldots$. It is assumed that an initial estimate γ_0 should be chosen before starting the iterative process—for example, using one of the ways proposed in [1, p. 186].

In our convergence analysis of the new method, we employ the notation used in Traub's book [1]: if $\{g_k\}$ and $\{h_k\}$ are null sequences and $g_k/h_k \to C$, where C is a nonzero constant, we shall write $g_k = O(h_k)$ or $g_k \sim Ch_k$.

We also use the concept of R-order of convergence introduced by Ortega and Rheinboldt [5]. Let $\{x_k\}$ be a sequence of approximations generated by an iterative method (IM). If this sequence converges to a zero α of f with the R-order $O_R((\mathrm{IM}), \alpha) \geq r$, we will write

$$\varepsilon_{k+1} \sim A_{k,r}\varepsilon_k^r, \tag{8}$$

where $A_{k,r}$ tends to the asymptotic error constant A_r of the iterative method (IM) when $k \to \infty$.

Introduce the abbreviations

$$\varepsilon_k = x_k - \alpha, \qquad \varepsilon_{k+1} = x_{k+1} - \alpha, \qquad w_k = x_k + \gamma_k f(x_k), \qquad \varepsilon_{k,w} = w_k - \alpha,$$

$$q_k = \gamma_k f'(\alpha), \qquad c_j = \frac{f^{(j)}(\alpha)}{j!f'(\alpha)} \ (p = 2, 3, \ldots).$$

Using Taylor's series about the root α, we obtain

$$f(x_k) = f'(\alpha)(\varepsilon_k + c_2\varepsilon_k^2 + c_3\varepsilon_k^3 + O(\varepsilon_k^4)) \tag{9}$$

and

$$\begin{aligned} f(x_k + \gamma_k f(x_k)) = f'(\alpha)\Big((1+q_k)\varepsilon_k + c_2(1+3q_k+q_k^2)\varepsilon_k^2 \\ + (2c_2^2q_k(1+q_k) + c_3(1+4q_k+3q_k^2+q_k^3))\varepsilon_k^3 + O(\varepsilon_k^4)\Big). \end{aligned} \tag{10}$$

In view of (9) and (10) we find from (1)

$$\begin{aligned}\varepsilon_{k+1} &= x_{k+1} - \alpha = \varepsilon_k - \frac{\gamma_k f(x_k)^2}{f(x_k + \gamma_k f(x_k)) - f(x_k)} \\ &= c_2(1+q_k)\varepsilon_k^2 + (-c_2^2(2+2q_k+q_k^2) + c_3(2+3q_k+q_k^2))\varepsilon_k^3 + O(\varepsilon_k^4). \end{aligned} \tag{11}$$

This error relation plays the key role in our study.

From the error relation (11) we observe that the choice $\gamma_k = -1/f'(\alpha)$ in (1) would provide cubic convergence. However, $f'(\alpha)$ is unknown in practice, so we content ourselves with some approximation $\bar{f}'(x_k)$ of $f'(\alpha)$. In this work we use the following approximation: $\bar{f}'(x_k) = N_2'(x_k)$, where $N_2(t) = N_2(t; x_k, x_{k-1}, w_{k-1})$ is Newton's interpolatory polynomial of second degree, set through three available approximations (nodes) $x_k,\ x_{k-1}$ and w_{k-1} to interpolate f. According to this formula, the self-correcting parameter γ_k is calculated as

$$\gamma_k = -\frac{1}{N_2'(x_k)}. \tag{12}$$

Combining (1) and (12), we construct the following derivative free method with memory of Steffensen's type:

$$\begin{cases} \gamma_0 \ \text{ is given,} \quad \gamma_k = -\dfrac{1}{N_2'(x_k)} \ \text{ for } k \geq 1, \\ x_{k+1} = x_k - \dfrac{\gamma_k f(x_k)^2}{f(x_k+\gamma_k f(x_k)) - f(x_k)}. \end{cases} \qquad (k = 0, 1, \ldots) \tag{13}$$

Lemma 1. $N_2'(x_k) \sim f'(\alpha)(1 - c_3\varepsilon_{k-1}\varepsilon_{k-1,w})$.

Proof. Using divided differences, we find

$$\begin{aligned} N_2'(x_k) &= \left[\frac{d}{dt}N_2(t)\right]_{t=x_k} \\ &= \left[\frac{d}{dt}(f(x_k) + f[x_k, x_{k-1}](t-x_k) + f[x_k, x_{k-1}, w_{k-1}](t-x_k)(t-x_{k-1}))\right]_{t=x_k} \\ &= f[x_k, x_{k-1}] + f[x_k, x_{k-1}, w_{k-1}](x_k - x_{k-1}) \\ &= f[x_k, w_{k-1}] + f[x_k, x_{k-1}] - f[x_{k-1}, w_{k-1}]. \end{aligned} \tag{14}$$

The estimation of $N_2'(x_k)$ is obtained using (9) and (14) as follows:

$$\begin{aligned} N_2'(x_k) &= f[x_k, x_{k-1}] + f[x_k, w_{k-1}] - f[w_{k-1}, x_{k-1}] \\ &= \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} + \frac{f(x_k) - f(w_{k-1})}{x_k - w_{k-1}} - \frac{f(w_{k-1}) - f(x_{k-1})}{w_{k-1} - x_{k-1}} \\ &= \frac{f(x_k) - f(x_{k-1})}{\varepsilon_k - \varepsilon_{k-1}} + \frac{f(x_k) - f(w_{k-1})}{\varepsilon_k - \varepsilon_{k-1,w}} - \frac{f(w_{k-1}) - f(x_{k-1})}{\varepsilon_{k-1,w} - \varepsilon_{k-1}} \\ &= f'(\alpha)\left[\frac{\varepsilon_k - \varepsilon_{k-1} + c_2(\varepsilon_k^2 - \varepsilon_{k-1}^2) + c_3(\varepsilon_k^3 - \varepsilon_{k-1}^3) + \ldots}{\varepsilon_k - \varepsilon_{k-1}}\right. \\ &\quad + \frac{\varepsilon_k - \varepsilon_{k-1,w} + c_2(\varepsilon_k^2 - \varepsilon_{k-1,w}^2) + c_3(\varepsilon_k^3 - \varepsilon_{k-1,w}^3) + \ldots}{\varepsilon_k - \varepsilon_{k-1,w}} \\ &\quad \left. - \frac{\varepsilon_{k-1,w} - \varepsilon_{k-1} + c_2(\varepsilon_{k-1,w}^2 - \varepsilon_{k-1}^2) + c_3(\varepsilon_{k-1,w}^3 - \varepsilon_{k-1}^3) + \ldots}{\varepsilon_{k-1,w} - \varepsilon_{k-1}}\right] \\ &= f'(\alpha)(1 + 2c_2\varepsilon_k + c_3(2\varepsilon_k^2 - \varepsilon_{k-1}\varepsilon_{k-1,w} + \varepsilon_k\varepsilon_{k-1} + \varepsilon_k\varepsilon_{k-1,w}) + \ldots) \\ &\sim f'(\alpha)(1 - c_3\varepsilon_{k-1}\varepsilon_{k-1,w}). \quad \square \end{aligned}$$

Now we state the following convergence theorem.

Theorem 1. *If an initial approximation x_0 is sufficiently close to a zero α of f, then the R-order of convergence of the two-point method (13) is at least 3.*

Proof. According to (9) we have

$$\varepsilon_{k,w} = w_k - \alpha = \varepsilon_k + \gamma_k f'(\alpha)(\varepsilon_k + c_2\varepsilon_k^2 + O(\varepsilon_k^3)) \sim (1 + \gamma_k f'(\alpha))\varepsilon_k, \tag{15}$$

while the error relation (11) gives

$$\varepsilon_{k+1} \sim (1 + \gamma_k f'(\alpha))\varepsilon_k^2. \tag{16}$$

Suppose that the R-orders of the sequences x_k and w_k are r and p, respectively. In view of (8) we have

$$\varepsilon_{k+1} \sim A_{k,r}\varepsilon_k^r \sim A_{k,r}(A_{k-1,r}\varepsilon_{k-1}^r)^r \sim A_{k,r}A_{k-1,r}^r\varepsilon_{k-1}^{r^2} \tag{17}$$

and

$$\varepsilon_{k,w} \sim A_{k,p}\varepsilon_k^p \sim A_{k,p}(A_{k-1,r}\varepsilon_{k-1}^r)^p \sim A_{k,p}A_{k-1,r}^p\varepsilon_{k-1}^{rp}. \tag{18}$$

By virtue of Lemma 1 we have

$$1 + \gamma_k f'(x_k) \sim c_3\varepsilon_{k-1}\varepsilon_{k-1,w}. \tag{19}$$

Combining (16)–(19) yields

$$\varepsilon_{k+1} \sim (1 + \gamma_k f'(\alpha))\varepsilon_k^2 \sim c_3\varepsilon_{k-1}\varepsilon_{k-1,w}\varepsilon_k^2 \sim c_3A_{k-1,p}A_{k-1,r}^2\varepsilon_{k-1}^{2r+p+1}. \tag{20}$$

Similarly, by (15) and (17)–(19),

$$\varepsilon_{w,k} \sim (1 + \gamma_k f'(\alpha))\varepsilon_k \sim c_3\varepsilon_{k-1}\varepsilon_{w,k-1}\varepsilon_k \sim c_3A_{k-1,p}A_{k-1,r}\varepsilon_{k-1}^{r+p+1}. \tag{21}$$

Equating exponents of the error ε_{k-1} in pairs of relations (17) $\wedge$ (20) and (18) $\wedge$ (21), we come to the system of equations

$$r^2 - 2r - p - 1 = 0,$$
$$rp - r - p - 1 = 0.$$

A positive solution is given by $r = 3$ and $p = 2$, which means that the R-order of convergence of the two-point method (13) is at least 3. $\square$

3. Numerical results

We have tested the new method (13) using the computational software package *Mathematica* with multiple-precision arithmetic. For comparison purposes, we have also tested the Steffensen-like method (1), Traub's method (7), the Newton-secant method (4) and three one-point methods displayed below.

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad \text{(Newton's method, order 2)}, \tag{22}$$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}\left(1 - \frac{f(x_k)f''(x_k)}{2f'(x_k)^2}\right)^{-1} \quad \text{(Halley's method, order 3)}, \tag{23}$$

$$x_{k+1} = x_k - \frac{f(x_k)}{\sqrt{f'(x_k)^2 - f(x_k)f''(x_k)}} \quad \text{(Ostrowski's method, order 3)}. \tag{24}$$

The errors $|x_k - \alpha|$ of approximations to the zeros are given in Tables 1 and 2, where $A(-h)$ denotes $A \times 10^{-h}$. These tables include the values of the computational order of convergence r_c calculated by using the formula

$$r_c = \frac{\log|f(x_k)/f(x_{k-1})|}{\log|f(x_{k-1})/f(x_{k-2})|}, \tag{25}$$

taking into consideration the last three approximations in the iterative process. For a demonstration, we have chosen two test functions:

$$f_1(x) = (x-2)\left(\frac{5}{x^2} + \frac{1}{5x} - 4x - x^5\right)e^{x^2-2x+\frac{1}{x^3}}, \qquad x_0 = 2.2, \qquad a = 2,$$

$$f_2(x) = x\log(1 + x\sin x) + e^{x\cos x + x^2 - 1}\sin\pi x, \qquad x_0 = 0.5, \qquad \alpha = 0.$$

Both examples have dealt with $\gamma_0 = 0.01$.

From Tables 1 and 2 and a number of numerical examples we can conclude that the proposed method (13) is certainly better than quadratically convergent methods, such as (1) and (22). The method (13) is competitive with third-order

Table 1
Errors of approximation to the zero of f_1.

Methods	$\|x_1-\alpha\|$	$\|x_2-\alpha\|$	$\|x_3-\alpha\|$	$\|x_4-\alpha\|$	r_c(25)
Steffensen-like IM (1)	1.13(−3)	2.88(−6)	1.88(−11)	7.97(−22)	1.9999
Traub's IM (7)	1.13(−3)	2.90(−6)	1.53(−13)	1.10(−30)	2.3559
Newton's IM (22)	9.29(−2)	2.59(−2)	2.47(−3)	2.42(−5)	1.8935
Halley's IM (23)	3.12(−2)	2.00(−4)	5.80(−11)	1.42(−29)	2.9998
Ostrowski's IM (24)	diverge				
Newton-secant IM (4)	4.81(−2)	1.39(−3)	4.31(−8)	1.29(−21)	2.9977
New method (13)	1.13(−3)	1.21(−8)	1.28(−23)	1.54(−68)	3.0000

Table 2
Errors of approximation to the zero of f_2.

Methods	$\|x_1-\alpha\|$	$\|x_2-\alpha\|$	$\|x_3-\alpha\|$	$\|x_4-\alpha\|$	r_c(25)
Steffensen-like IM (1)	2.60(−2)	6.71(−4)	4.55(−7)	2.10(−13)	1.9998
Traub's IM (7)	2.60(−2)	2.04(−4)	1.07(−9)	2.32(−22)	2.3981
Newton's IM (22)	2.94(−2)	8.49(−4)	7.20(−7)	5.18(−13)	1.9997
Halley's IM (23)	1.63(−1)	3.46(−3)	1.22(−8)	5.14(−25)	3.0032
Ostrowski's IM (24)	1.48(−1)	1.37(−3)	5.50(−10)	3.67(−29)	2.9980
Newton-secant IM (4)	8.95(−3)	7.11(−7)	3.59(−19)	4.63(−56)	2.9999
New method (13)	2.60(−2)	1.86(−4)	2.11(−12)	2.62(−36)	3.0089

methods considering the accuracy of the approximations produced, but it attains such accuracy using only two function evaluations per iteration. A better insight into the computational efficiency can be obtained by comparing efficiency indices.

Using (5) we calculate the efficiency indices

$$E(1)=E(22)\approx 1.414,\qquad E(4)=E(23)=E(24)\approx 1.442,\qquad E(7)\approx 1.554,\qquad E(13)\approx 1.732.$$

We observe that the proposed method (13) is significantly more efficient than the methods considered. Moreover, the efficiency index of (13) is even higher than that of optimal two-point methods ($=4^{1/3}\approx 1.587$) and optimal three-point methods ($=8^{1/4}\approx 1.682$); see, e.g., [6,7]. Having in mind that very fast iterative methods produce very accurate approximations not required for real-life problems, it turns out that the new method (13) is of important practical interest.

We end this work with an analogous method for solving systems of nonlinear equations. Introduce an n-dimensional divided difference $[x, y; f]$ as an $n\times n$ matrix with elements

$$[x,y;f]_{i,j}=\frac{f_i(x^{(1)},\ldots,x^{(j)},y^{(j+1)},\ldots,y^{(n)})-f_i(x^{(1)},\ldots,x^{(j-1)},y^{(j)},\ldots,y^{(n)})}{x^{(j)}-y^{(j)}},$$

where $x=(x^{(1)},\ldots,x^{(n)})$ and $f=(f_1,\ldots,f_n)$ are vectors (see [8, p. 20]). Having in mind that $N_2'(x_k)$ is a matrix $N_2'(x_k)=[x_k,w_{k-1};f]+[x_k,x_{k-1};f]-[x_{k-1},w_{k-1};f]$ (see the proof of Lemma 1), starting from (13) we can construct the following analogous method for solving systems of nonlinear equations:

$$\begin{cases}\gamma_0 \ \text{ is given}, \quad \gamma_k=-[N_2'(x_k)]^{-1} \ \text{ for } k\ge 1,\\ x_{k+1}=x_k-[x_k,x_k+\gamma_k f(x_k);f]^{-1}f(x_k).\end{cases}\qquad (k=0,1,\ldots) \tag{26}$$

Unfortunately, the two-step method (26) does not keep the advantage of low computational cost of the scalar method (13) since the calculation of γ_k in the iterative scheme (26) requires the inverse of the sum of three matrices. A simple analysis shows that the method (26) needs $8n^2+n$ function evaluations (of n variables) plus two matrix inversions, while some existing two-step methods of the form

$$\begin{cases}y_k=x_k-[g_1(x_k),g_2(x_k);f]^{-1}f(x_k),\\ x_{k+1}=y_k-[g_1(x_k),g_2(x_k);f]^{-1}f(y_k)\end{cases}\qquad (k=0,1,\ldots) \tag{27}$$

(presented, e.g., in [8–10]) require only $2n^2+2n$ function evaluations plus one matrix inversion. Because of this low computational efficiency, we did not study the method (26) in detail, restricting our research to the scalar method (13), which is more efficient than scalar versions of the methods of the form (27).

Acknowledgment

This work was supported by the Serbian Ministry of Science under grant 174022.

References

[1] J.F. Traub, Iterative Methods for the Solution of Equations, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.
[2] I.F. Steffensen, Remarks on iteration, Skand. Aktuarietidskr. 16 (1933) 64–72.
[3] J. Džunić, M.S. Petković, L.D. Petković, Three-point methods with and without memory for solving nonlinear equations, Appl. Math. Comput. 218 (2012) 4917–4927.
[4] M.S. Petković, J. Džunić, L.D. Petković, A family of two-point methods with memory for solving nonlinear equations, Appl. Anal. Discrete Math. 5 (2011) 298–317.
[5] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.
[6] W. Bi, Q. Wu, H. Ren, A new family of eight-order iterative methods for solving nonlinear equations, Appl. Math. Comput. 214 (2009) 236–245.
[7] C. Chun, Some fourth-order iterative methods for solving nonlinear equations, Appl. Math. Comput. 195 (2008) 454–459.
[8] S. Amat, S. Busquier, Convergence and numerical analysis of a family of two-step Steffensen's methods, Comput. Math. Appl. 49 (2005) 13–22.
[9] S. Amat, S. Busquier, V.F. Candela, A class of quasi-Newton generalized Steffensen methods on Banach spaces, J. Comput. Appl. Math. 149 (2003) 397–406.
[10] I.K. Argyros, An error analysis for the Steffensen method under Zabrejko–Nguyen-type assumptions, Revue d'Analysis Numerique et de Theorie de l'Approximation 25 (1996) 11–22.