Generalized Newton-Raphson method (GNR)

Nikola Sandrić
Department of Mathematics
Faculty of Civil Engineering, University of Zagreb, Zagreb, Croatia
Email: nsandric@grad.hr

Radimir Viher
Department of Mathematics
Faculty of Civil Engineering, University of Zagreb, Zagreb, Croatia
Email: viher@grad.hr

September 3, 2012

Abstract

We present a new method for solving a non-linear equation $f(x) = 0$. The presented method is quadratically convergent, it converges faster than the classical Newton-Raphson method and the Newton-Raphson method appears as the limiting case of the presented method.

Keywords and phrases: Newton-Raphson method, generalized Newton-Raphson method, Aitken's $\triangle^2$-method, Steffensen's function

1 Introduction

Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be an arbitrary function. *Zero* of the function $f(x)$ is a point $\xi \in \mathbb{R}$ such that $f(\xi) = 0$. The problem of finding zeros of a given function is complex and analytically very hard solvable. The aim of this paper is to find a numerical (approximative) method for finding zeros of a given function. This problem has been already studied in the literature (see, e.g., [SB93]). The most well known method which considers this type of problem is the *Newton-Raphson method* (NR): Let $f : (a, b) \longrightarrow \mathbb{R}$ be twice continuously differentiable function with a simple zero $\xi \in (a, b)$ (i.e., $f'(\xi) \neq 0$). Then for an arbitrary $x_0 \in (a, b)$ the following iterative sequence

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \ldots,$$

converges to ξ and this convergence is quadratic. Recall that a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ converges *quadratically* to ξ if there exist $n_0 \in \mathbb{N}$ and $C \geq 0$ such that

$$|x_{n+1} - \xi| \leq C|x_n - \xi|^2,$$

for all $n \geq n_0$.

In this paper, we present a method which is also quadratically convergent, it converges faster than (NR) and (NR) appears as the limiting case of the presented method.

First, recall that a mapping $\Phi : \mathbb{R} \longrightarrow \mathbb{R}$ is called a *contraction mapping* if there exists a nonnegative real number $q < 1$, such that $|\Phi(y) - \Phi(x)| \leq q|y - x|$ for all $x, y \in \mathbb{R}$. Further, the famous Banach fixed point theorem states that every contraction mapping $\Phi(x)$ admits one and only one fixed point ξ in $\mathbb{R}$. Furthermore, this fixed point can be found as follows. Start with an arbitrary element x_0 in $\mathbb{R}$ and define an iterative sequence by $x_{n+1} := \Phi(x_n)$, for $n = 0, 1, \ldots$. Then, this sequence converges and its limit is ξ.

Now, we present our method for finding zeros of a given function. Let $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ be a convergent sequence with limit ξ. The *Aitken's accelerating convergence* $\triangle^2$*-method* (see [SB93]) is based on introducing a new sequence $\{\bar{x}_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ defined by the formula

$$\bar{x}_n = x_n - \frac{(\triangle x_n)^2}{\triangle^2 x_n}, \quad n \in \mathbb{N},$$

where $\triangle x_n = x_{n+1} - x_n$ and $\triangle^2 x_n = \triangle(\triangle x_n)$. More precisely, if there exists $\mu \in (0, 1)$ such that

$$\lim_{n\longrightarrow\infty} \frac{|x_{n+1} - \xi|}{|x_n - \xi|} = \mu$$

and if $(x_n - \xi)(x_{n+1} - \xi) > 0$ for all $n \in \mathbb{N}$ large enough, then

$$\lim_{n\to\infty} \frac{\bar{x}_n - \xi}{x_n - \xi} = 0.$$

Hence, $\{\bar{x}_n\}_{n\in\mathbb{N}}$ converges faster than $\{x_n\}_{n\in\mathbb{N}}$ to ξ. Now, let $\Phi : \mathbb{R} \longrightarrow \mathbb{R}$ be a contraction mapping with fixed point ξ. By the Banach fixed point theorem, for an arbitrary initial point $x_0 \in \mathbb{R}$ the iterative sequence $x_{n+1} := \Phi(x_n)$ converges to ξ. Further, by using Aitken's idea to get faster convergence of the sequence $\{x_n\}_{n\in\mathbb{N}}$, Steffensen introduced a function $\Psi(x)$ defined by the formula

$$\Psi(x) := \frac{x\Phi(\Phi(x)) - (\Phi(x))^2}{\Phi(\Phi(x)) - 2\Phi(x) + x} \tag{1.1}$$

(see [SB93]). Under the assumptions that $\Phi(x)$ is continuously differentiable and $\Phi'(x) \neq 1$, it can be shown that ξ is the unique fixed point of $\Psi(x)$. Further, if $\Phi(x)$ is twice continuously differentiable, it can be shown that for an arbitrary initial point $x_0 \in \mathbb{R}$ the iterative sequence $x_{n+1} := \Psi(x_n)$ converges to ξ quadratically.

Now, let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be twice continuously differentiable function with simple zero ξ. Let us assume that (a, b) is a neighborhood around ξ and that $f'(x)$ does not change the sign on $[a, b]$. Without loss of generality, assume that

$$0 < m \leq f'(x) \leq M \quad \text{for all} \quad x \in [a, b].$$

Now, let $0 < k \leq \frac{1}{M}$ be arbitrary and let us define function $\Phi(x)$ by the formula $\Phi(x) := x - kf(x)$. Then we have

$$0 \leq 1 - kM \leq \Phi'(x) = 1 - kf'(x) \leq 1 - km < 1,$$

i.e., $\Phi(x)$ is a contraction. Due to this, from the Banach fixed point theorem, for an arbitrary initial point $x_0 \in \mathbb{R}$ the iterative sequence

$$x_{n+1} = x_n - kf(x_n)$$

converges to ξ. Unfortunately, in general, this convergence is not quadratic (e.g., take $f(x) = 2x$). To get a quadratic convergence, we insert $\Phi(x)$ in (1.1) and we get

$$\Psi(x) = x - \frac{k(f(x))^2}{f(x) - f(x - kf(x))}.$$

Note that $f'(\xi) \neq 0$ implies $\Phi'(\xi) \neq 1$. Hence, ξ is the unique fixed point of $\Psi(x)$. The new recursion is given by

$$x_{n+1} := x_n - \frac{k(f(x_n))^2}{f(x_n) - f(x_n - kf(x_n))}, \quad n = 0, 1, \ldots, \tag{1.2}$$

and, as we discussed before, for an arbitrary initial point $x_0 \in \mathbb{R}$ it converges quadratically to ξ. We call the iterative recursion (1.2) the *generalized Newton-Raphson method (GNR)*. Observe that if the constant k converges to zero, we get

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \ldots,$$

and this is nothing else than (NR).

2 Main result

In this section, we show that by increasing the value of k, the convergence of (GNR) accelerates. According to this, we get that (GNR) converges faster than (NR). First, we need the following technical lemma.

Lemma 2.1. *Let $f(x)$ be twice continuously differentiable function with $f''(x) > 0$ (or $f''(x) < 0$) on (a, b). Then there exists a unique function $c : (a, b) \longrightarrow (a, b)$ such that*

$$f(b) - f(x) = f'(c(x))(b - x). \tag{2.1}$$

Moreover, the function $c(x)$ is continuously differentiable on (a, b) and

$$c'(x) = \frac{f'(c(x)) - f'(x)}{(b - x)f''(c(x))}. \tag{2.2}$$

Proof. Without loss of generality, we consider the case when $f''(x) > 0$. Uniqueness of $c(x)$ follows from Lagrange's mean value theorem and from the fact that the function $f'(x)$ is strictly increasing on (a, b). The existence of $c'(x)$ and (2.2) follows from the inverse function theorem and

$$c(x) = f'^{-1}\left(\frac{f(b) - f(x)}{b - x}\right).$$

□

Theorem 2.2. *Let $f(x)$ be twice continuously differentiable function with simple zero ξ. Further, let us assume that $f''(x) > 0$ on $(\xi, \xi + l)$, for some $l > 0$. Then for any two constants k_1 and k_2 which satisfy*

$$0 < k_1 \leq k_2 \leq \frac{1}{\max\limits_{x \in (\xi, \xi + l)} f'(x)} = \frac{1}{f'(\xi + l)} \tag{2.3}$$

and for any initial approximations $x_0^{(1)} = x_0^{(2)} = x_0 \in (\xi, \xi + l)$, the (GNR) iterative sequence $\{x_n^{(2)}\}_{n \in \mathbb{N}}$ with constant k_2 is always faster than the (GNR) iterative sequence $\{x_n^{(1)}\}_{n \in \mathbb{N}}$ with constant k_1.

Proof. Let k satisfy condition (2.3). First, we show that

$$\xi < x - kf(x) \tag{2.4}$$

holds for all $x \in (\xi, \xi + l)$. Indeed, (2.4) holds if and only if

$$k\frac{f(x) - f(\xi)}{x - \xi} < 1, \tag{2.5}$$

and (2.5) holds because of condition (2.3) and the Lagrange's mean value theorem.

Now, we show that if the constant k satisfies condition (2.3), then the function $g(x)$ defined by

$$g(x) := x - \frac{k(f(x))^2}{f(x) - f(x - kf(x))}$$

is strictly increasing on segment $(\xi, \xi + l)$. We have

$$\begin{aligned} g'(x) &= 1 - \frac{f'(x) + 2f'(x)\frac{f(x)-f(x-kf(x))-f(x)}{f(x)} + f'(x - kf(x))(1 - kf'(x))}{k\left(\frac{f(x)-f(x-kf(x))}{kf(x)}\right)^2} \\ &= 1 - \frac{kf'(x)(2f'(c) + f'(x) - f'(x - kf(x)) - f'(x)) - f'(x)}{k(f'(c))^2} + \frac{f'(x - kf(x))}{k(f'(c))^2} \\ &= 1 - \frac{f'(x)(2f'(c) + f''(\bar{c})kf(x) - f'(x))}{(f'(c))^2} - \frac{f'(x) - f'(x - kf(x))}{k(f'(c))^2} \\ &= 1 - \frac{f'(x)(2f'(c) - f'(x)) - f''(\bar{c})f(x)(1 - kf'(x))}{(f'(c))^2}, \end{aligned} \tag{2.6}$$

where, by (2.4) and the Lagrange's mean value theorem, $c, \bar{c} \in (x - kf(x), x)$. Next, note that (2.6) is equivalent with

$$0 \le (f'(x) - f'(c))^2 + f''(\bar{c})f(x)(1 - kf'(x)), \tag{2.7}$$

and clearly (2.7) holds for all $x \in (\xi, \xi + l)$.

Now, we prove the claim of the theorem by induction. Let $n = 1$. We have

$$x_1^{(1)} = x_0 - \frac{k_1(f(x_0))^2}{f(x_0) - f(x_0 - k_1 f(x_0))} \quad \text{and} \quad x_1^{(2)} = x_0 - \frac{k_2(f(x_0))^2}{f(x_0) - f(x_0 - k_2 f(x_0))}. \tag{2.8}$$

Clearly, by Lagrange's mean value theorem, (2.8) can be rewritten in the form

$$x_1^{(1)} = x_0 - \frac{f(x_0)}{f'(c_0^{(1)})} \quad \text{and} \quad x_1^{(2)} = x_0 - \frac{f(x_0)}{f'(c_0^{(2)})}. \tag{2.9}$$

Note that, since $f''(x) > 0$, the constants $c_0^{(1)}$ and $c_0^{(2)}$ are unique. Next, from Lemma 2.1 and the fact that $x_0 - k_2 f(x_0) \le x_0 - k_1 f(x_0)$, we have

$$c_0^{(2)} \le c_0^{(1)}. \tag{2.10}$$

Thus, from (2.9) and (2.10) we have

$$x_1^{(1)} - x_1^{(2)} = \frac{f(x_0)(f'(c_0^{(1)}) - f'(c_0^{(2)}))}{f'(c_0^{(1)})f'(c_0^{(2)})} \ge 0. \tag{2.11}$$

Now, let $n \in \mathbb{N}$ be arbitrary and let us prove that $x_n^{(1)} - x_n^{(2)} \geq 0$ implies $x_{n+1}^{(1)} - x_{n+1}^{(2)} \geq 0$. Clearly, as in the case $n = 1$,

$$x_n^{(2)} - k_2 f(x_n^{(2)}) \leq x_n^{(2)} - k_1 f(x_n^{(2)})$$

implies $c_n^{(2)} \leq c_n^{(1)}$. Hence, we have

$$\begin{aligned} x_{n+1}^{(1)} - x_{n+1}^{(2)} &\geq x_n^{(2)} - \frac{k_1 (f(x_n^{(2)}))^2}{f(x_n^{(2)}) - f(x_n^{(2)} - k_1 f(x_n^{(2)}))} - \left(x_n^{(2)} - \frac{k_2 (f(x_n^{(2)}))^2}{f(x_n^{(2)}) - f(x_n^{(2)} - k_2 f(x_n^{(2)}))} \right) \\ &= \frac{f(x_n^{(2)})}{f'(c_n^{(2)})} - \frac{f(x_n^{(2)})}{f'(c_n^{(1)})} = \frac{f(x_n^{(2)})(f'(c_n^{(1)}) - f'(c_n^{(2)}))}{f'(c_n^{(1)}) f'(c_n^{(2)})} \geq 0, \end{aligned}$$

where the first inequality is a consequence of strictly increasing property of the function $g(x)$. $\square$

Remark 2.3. Clearly, according to Theorem 2.2 we choose k to be exactly $(f'(\xi + l))^{-1}$. Further, as we commented in the first section, (NR) is a special case of (GNR) when $k \longrightarrow 0$. Hence, (GNR) is faster than (NR).

In the previous theorem we only considered the case when $f(x)$, $f'(x)$ and $f''(x) > 0$ are positive on $(\xi, \xi + l)$. In what follows, we discuss other possible cases.

1. The first case in which $f(x)$, $f'(x)$ and $f''(x)$ are positive on $(\xi, \xi + l)$ (see Figure 1), was treated in Theorem 2.2.

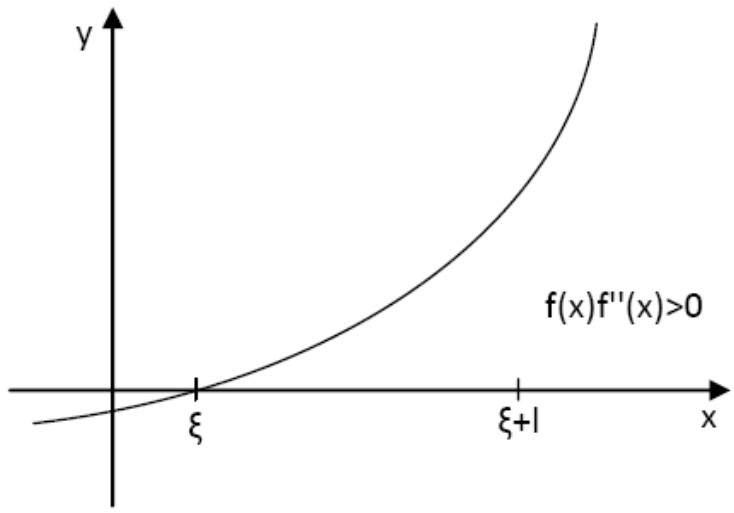

Figure 1: Case 1.

2. In the second case, $f(x)$ is positive, $f'(x)$ is negative and $f''(x)$ is positive on $(\xi - l, \xi)$ (see Figure 2). Then, under the assumption

$$\frac{1}{f'(\xi - l)} = \frac{1}{\min\limits_{x \in (\xi - l, \xi)} f'(x)} \leq k < 0,$$

Theorem 2.2 is also true. In this case, we choose k to be $(f'(\xi - l))^{-1}$.

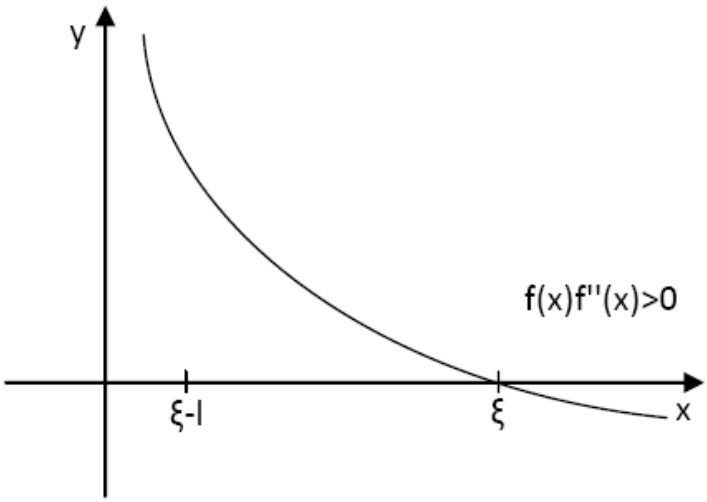

Figure 2: Case 2.

3. In the third case, $f(x)$ is negative, $f'(x)$ is positive and $f''(x)$ is negative on $(\xi - l, \xi)$. Similarly as in the first two cases, under the assumption

$$0 < k \leq \frac{1}{\max\limits_{x \in (\xi - l, \xi)} f'(x)} = \frac{1}{f'(\xi - l)}$$

Theorem 2.2 holds. In this case, we choose k to be $(f'(\xi - l))^{-1}$.

4. In the fourth case, $f(x)$, $f'(x)$ and $f''(x)$ are negative on $(\xi, \xi + l)$. In this case, under the assumption

$$\frac{1}{f'(\xi + l)} = \frac{1}{\min\limits_{x \in (\xi, \xi + l)} f'(x)} \leq k < 0$$

Theorem 2.2 holds true. In this case, we choose k to be $(f'(\xi + l))^{-1}$.

3 Examples

In this section, we give two examples and application of (GNR) method.

Example 3.1. *Let us demonstrate our method on the problem of finding square roots of* 29.

Solution. Clearly, we have to find zeros of the function $f(x) = x^2 - 29$. The (NR) recursion for this problem is given by

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{29}{x_n}\right), \quad n \in \mathbb{N}.$$

With the initial approximation $x_0 = 50$ and the iteration precision $\varepsilon = 10^{-9}$ (i.e., we iterate until $|x_{n+1} - x_n| < \varepsilon$), we get:

$x_0 = 50$
$x_1 = 25.29$
$x_2 = 13.2183491498$
$x_3 = 7.7061345534$
$x_4 = 5.7346850838$
$x_5 = 5.39581617$
$x_6 = 5.38517532$
$x_7 = 5.3851648071$

In the case of (GNR), we have to take the constant k such that

$$0 < k \leq \frac{1}{\max\limits_{x \in (\sqrt{29}, 50)} f'(x)} = 0.01.$$

We consider three different choices of such k and we get:

$k = 0.0001$	$k = 0.001$	$k = 0.01$
$x_0 = 50$	$x_0 = 50$	$x_0 = 50$
$x_1 = 25.2287903409$	$x_1 = 24.6639461083$	$x_1 = 17.1802364191$
$x_2 = 13.1746225681$	$x_2 = 12.780313516$	$x_2 = 8.7837054533$
$x_3 = 7.6849000217$	$x_3 = 7.4969469572$	$x_3 = 5.9653845646$
$x_4 = 5.7288844854$	$x_4 = 5.6792962793$	$x_4 = 5.4103182009$
$x_5 = 5.3954648792$	$x_5 = 5.3926992458$	$x_5 = 5.3852169789$
$x_6 = 5.385174628$	$x_6 = 5.3851700138$	$x_6 = 5.3851648073$
$x_7 = 5.3851648071$	$x_7 = 5.3851648071$	

Thus, as discussed in previous section, (GNR) is faster than (NR) and the numerical results confirm Theorem 2.2.

In order to find the negative root of 29, we proceed in the same way. For the initial approximation we take $x_0 = -50$. The constant k has to satisfy the following relation:

$$-0.01 = \frac{1}{\min\limits_{x \in (-50, -\sqrt{29})} f'(x)} \leq k < 0.$$

We consider three different cases of such k and we get:

$k = -0.0001$	$k = -0.001$	$k = -0.01$
$x_0 = -50$	$x_0 = -50$	$x_0 = -50$
$x_1 = -25.2287903409$	$x_1 = -24.6639461083$	$x_1 = -17.1802364191$
$x_2 = -13.1746225681$	$x_2 = -12.780313516$	$x_2 = -8.7837054533$
$x_3 = -7.6849000217$	$x_3 = -7.4969469572$	$x_3 = -5.9653845646$
$x_4 = -5.7288844854$	$x_4 = -5.6792962793$	$x_4 = -5.4103182009$
$x_5 = -5.3954648792$	$x_5 = -5.3926992458$	$x_5 = -5.3852169789$
$x_6 = -5.385174628$	$x_6 = -5.3851700138$	$x_6 = -5.3851648073$
$x_7 = -5.3851648071$	$x_7 = -5.3851648071$	

□

Example 3.2. *Problem is to find a minimum of the function*

$$f(x) = (e^x - x)^2 + (x^2 - \cos x)^2$$

(with the iteration precision $\varepsilon = 10^{-9}$).

Solution. It is obvious that minimum exists since the function $f(x)$ is sum of two squares. Further, since

$$f'(x) = 2(e^x - x)(e^x - 1) + 2(x^2 - \cos x)(2x + \sin x),$$

it is easy to see that one zero of $f'(x)$ is $x_M = 0$. But, since $f''(0) = -4$, x_M is local maximum. Hence, local minimums are on the left and on the right to the origin. To find negative minimum, for the initial approximation we take $x_0 = -1$ and $k = 0.1$. After four (GNR) iterations we get $x_4 = -0.690308229$. Analogously, to find positive minimum, for the initial approximation we take $\bar{x}_0 = 1$ and $\bar{k} = 0.1$. After four (GNR) iterations we get $\bar{x}_4 = 0.5566918834$. □

References

[SB93] J. Stoer and R. Bulirsch. *Introduction to numerical analysis.* Springer-Verlag, New York, second edition, 1993.