sciendo

DOI: 10.2478/awutm-2019-0003

Analele Universităţii de Vest,
Timişoara
Seria Matematică – Informatică
LVII, 1, (2019), 13– 22

On the local convergence of the Modified Newton method

Ştefan Măruşter

Abstract. The aim of this paper is to investigate the local convergence of the Modified Newton method, i.e. the classical Newton method in which the first derivative is re-evaluated periodically after m steps. The convergence order is shown to be $m+1$. A new algorithm is proposed for the estimation the convergence radius of the method. We propose also a threshold for the number of steps after which is recommended to re-evaluate the first derivative in the Modified Newton method.

.

AMS Subject Classification (2000). 45G10; 47H17; 65J15; 65G99
Keywords. modified Newton method; local convergence; radius of convergence

1 Introduction

We investigate the local convergence of the Modified Newton method, i.e. the classical Newton method in which the first derivative is re-evaluated periodically after m steps. If x denotes the current iteration, then the iteration function T for the Modified Newton is defined by

$$\begin{aligned} &y_{k+1} = y_k - F'(x)^{-1}F(y_k),\ \ k = 1, ..., m-1,\ \ y_1 = x, \\ &T(x) = x - F'(x)^{-1}\textstyle\sum_{k=1}^{m} F(y_k). \end{aligned} \tag{1.1}$$

The iteration (1.1) can be interpreted also as a Picard iteration with the iteration function T.

The particular case $m = 2$ was considered by Potra and Ptak [6]. Using non-discrete induction, they proved the order three of convergence and gave sharp a priori and a posteriori error for this particular case. Often it is called "Potra-Ptak" method [7,8]. In the case of a single equation, Potra-Ptak method was considered by Traub [9] (1982). Ortega and Rheinboldt [5] proved order three of convergence for Potra-Ptak method in n-dimensional spaces (Theorem 10.2.4, [5]). Note that Potra-Ptak method is a particular case of a multipoint iterative process with order three of convergence considered by Ezquerro and Hernandez [2].

Recently, Hernandes and Romero [3] gave the following algorithm (formula) to estimate the local convergence radius for the Ezquerro-Hernandez method. Suppose that x^* is a solution of the equation $F(x) = 0$, there exists $F'(x^*)^{-1}$, $\|F'(x^*)^{-1}\| \leq \beta$, and F' is k-Lipschitz continuous on some $B(x^*, r_0) = \{x : \|x - x^*\| \leq r_0\}$. Let $\tilde{r} = \min\{r_0, r\}$, where $r = \zeta_0/[(1+\zeta_0)\beta k]$ and ζ_0 is the positive real root of a polynomial equation of degree three (in the particular case of Potra-Ptak this equation is $t^3 + 4t^2 - 8 = 0$). Then $\tilde{r}$ estimates the local radius of convergence.

In [1] Catinas proposes a simple and elegant formula to estimate the radius of convergence for the general Picard iteration and the algorithm presumptively gives a sharp value. More precisely, let $T : D \subset \mathcal{R}^m \to D$ be a nonlinear mapping and x^* a fixed point of T. Suppose that T is differentiable on some ball centred in x^*, $B(x^*, r_1)$, and the derivative of T satisfies

$$\begin{aligned}
&\|T'(x^*)\| \leq q < 1, \\
&\|T'(x) - T'(y)\| \leq k\|x - y\|^p, \ \forall x \in B(x^*, r_1).
\end{aligned}$$

Define

$$r_2 = \left(\frac{(1+p)(1-q)}{k}\right)^{\frac{1}{p}},$$

then $r = \min\{r_1, r_2\}$ is an estimation of local convergence radius.

In [4] a new algorithm (formula) of the radius of convergence of Modified Newton method was recently proposed. The algorithm was motivated in [4] by heuristic arguments and numerous examples for different iterative methods show that the proposed algorithm provides satisfactory estimations of the convergence radius.

Suppose that the mapping $F : C \to \mathcal{H}$, where C is an open convex subset of the Hilbert space $\mathcal{H}$, is Fréchet differentiable and satisfies

1. There exists $F'(x)^{-1}$ and $\|F'(x)^{-1}(x)\| \leq \beta$ for all $x \in C$;

2. F' is L-Lipschitz continuous on C;

then the formula for the radius of convergence proposed in [4] is $r \leq \alpha_m/(\beta L)$. The number α_m is the smallest positive solution of the polynomial equation $g_m(x) - \eta = 0$, where $0 < \eta < 1$ and g_m is defined by

$$\begin{aligned} f_1(y) &= y, \\ f_k(y) &= y\left(1 + \tfrac{1}{2}\prod_{j=1}^{k-1} f_j(y)\right), \; k = 2, 3, ..., \\ g_m(y) &= \prod_{k=1}^{m} f_k(y). \end{aligned} \tag{1.2}$$

The convergence of the Modified Newton method is obtained with the help of the inequality

$$\|T(x) - p\| \leq \eta\|x - p\|, \; \forall x \in B(p, r),$$

where p is a solution of the equation $F(x) = 0$ (or, equivalent, a fixed point of T).

In this paper we complete the result of [4] by providing a complete proof of the local convergence of the method (1.1) and by developing the investigation of the corresponding radius of convergence. The formula giving the estimation of convergence radius of (1.1) is proved to be

$$r \leq \frac{\alpha}{\beta L},$$

where α is the unique solution in $(0, 1)$ of the equation $t^3 + 2t^2 - 2 = 0$ ($\alpha = \sqrt{3} - 1$).

2 Preliminaries

The sequence of polynomials $\{f_k\}$ and the polynomial g_m defined by (1.2) can also be defined by the following recurrence formula

$$\begin{aligned} &f_1(y) = y, \; f_2(y) = y(1 + \tfrac{1}{2}y), \\ &f_k(y) = f_{k-1}^2(y) - yf_{k-1}(y) + y, \; k = 3, 4, ..., \end{aligned} \tag{2.1}$$

and

$$g_m(y) = 2\left(\frac{f_m(y)}{y} - 1\right) f_m(y).$$

It is easy to show that the two sequences (defined by (1.2) and (2.1)) and the polynomial g_m are identical.

Lemma 2.1. *Suppose* $0 < y < \alpha = \sqrt{3} - 1$. *Then*
(a) $f_k(y) < 1$ *and* $f_k(\alpha) = 1,\ k = 2, 3, ...$;
(b) $g_m(y) < \alpha$ *and* $g_m(\alpha) = \alpha$.

Proof. Both statements can be obtained very easily by induction on k and m, respectively.

(a) For $k = 2$, $f_2(y) = \frac{1}{2}y^2 + y$ and $y < \alpha$ implies $f_2(y) < 1$ and $f_2(\alpha) = 1$.

Suppose that $f_k(y) < 1$. We must prove that $f_{k+1}(y) = f_k(y)^2 - yf_k(y) + y < 1$. The quadratic polynomial in $f_k(y)$, $P(f_k(y)) = f_k(y)^2 - yf_k(y) + y - 1$ has the zeros $y - 1$ and 1. Thus $P(f_k(y)) < 0$. If $f_k(\alpha) = 1$, then $f_{k+1}(\alpha) = f_k(\alpha)^2 - \alpha f_k(\alpha) + \alpha = 1$.

(b) For $m = 2$,

$$g_2(y) = 2\left(\frac{f_2(y)}{y} - 1\right) f_2(y) = \frac{1}{2}y^3 + y^2,$$

and $y < \alpha$ implies $g_2(y) < \alpha$.

Suppose that for $y < \alpha$, $g_m(y) = 2[f_m(y)^2 - yf_m(y)]/y < \alpha$. We must prove that $g_{m+1}(y) = 2(f_{m+1}(y)/y - 1)f_{m+1}(y) < \alpha$. We have

$$g_{m+1}(y) = 2\left(\frac{f_m(y)^2 - yf_m(y) + y}{y} - 1\right) f_{m+1}(y)$$

$$= \frac{2}{y}[f_m(y)^2 - yf_m(y)]f_{m+1}(y) < \alpha f_{m+1}(y) < \alpha.$$

Using the first definition of f_k, formulas (1.2) and (a), we have

$$g_m(\alpha) = \prod_{k=1}^{m} f_k(\alpha) = f_1(\alpha) = \alpha.$$

□

Am important consequence of this Lemma is that the equation $g_m(y) - \eta = 0$ has α_m as the smallest zero in $(0, 1)$ for any η with $0 < \eta < \alpha$.

Lemma 2.2. *Let* d *and* $d_k,\ k = 1, ..., m$ *be linear mappings and suppose that* d *is invertible. Then*

$$d^{-1}\left(d_1 + \sum_{k=2}^{m} d_k \prod_{j=1}^{k-1} d^{-1}(d - d_{k-j})\right) = I - \prod_{j=0}^{m-1} d^{-1}(d - d_{m-j}).$$

The proof can be obtained easily by induction on m.

3 Local convergence and radius of convergence

Let $\mathcal{H}$ be a real Hilbert space with scalar product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$ and C an open subset of $\mathcal{H}$. Let $F : C \to H$ be a Fréchet differentiable mapping and suppose that the set of solutions of the equation $F(x) = 0$ (or the set of fixed points of T defined by (1.1)) is nonempty.

Let $\{r_k\}$, $k = 1, 2, ..., m$ be a numerical sequence defined recursively by

$$r_{k+1} = \widetilde{\alpha} r_k \left(1 + \frac{r_k}{2\widetilde{r}}\right),\ r_1 = \widetilde{r}, \tag{3.1}$$

where $\widetilde{\alpha} > 2/3$ and $\widetilde{r} > 0$. Then $\{r_k\}$ is strictly increasing (this can be easily shown by induction on k).

Theorem 3.1. *Suppose that there exists $F'(x)^{-1}$, $\|F'(x)^{-1}\| \le \beta$, $\forall x \in C$ and that F' is L-Lipschitz continuous on C. Let r be such that $0 < r \le \alpha/(\beta L)$, $\alpha = \sqrt{3} - 1$. Let p be a solution of $F(x) = 0$ and suppose that $B(p, r_m) \subset C$, where r_m is defined by (3.1) for $\widetilde{\alpha} = \alpha$ and $\widetilde{r} = r$. Then the sequence $\{x_n\}$ given by the Modified Newton method with starting point $x_0 \in B(p, r)$, remains in $B(p, r)$ and converges to the unique solution p in $B(p, r)$. The rate of convergence is at least $m + 1$.*

Proof. For any $x \in B(p, r)$, we have $\|y_k - p\| \le r_k$, $k = 1, 2, .., m$. Indeed, for $k = 1$, $\|y_1 - p\| = \|x - p\| \le r = r_1$. Using the definition of y_{k+1} and the Mean Value Theorem, we get

$$\|y_{k+1} - p\| \le \beta L r \|y_k - p\| \left(1 + \frac{\|y_k - p\|}{2r}\right).$$

Supposing that $\|y_k - p\| \le r_k$, we have

$$\|y_{k+1} - p\| \le \alpha r_k (1 + \frac{r_k}{2r}) = r_{k+1}.$$

We can conclude that $y_k \in C$, $k = 1, ..., m$ and the sequence $\{x_n\}$ is well defined.

From the definition of y_k and Mean Value Theorem we can obtain

$$y_k - p = \left(\prod_{j=1}^{k-1} F'(x)^{-1}(F'(x) - \Delta_{k-j})\right)(x - p),\ k = 2, ..., m, \tag{3.2}$$

where

$$\Delta_j = \int_0^1 F'(p + t(y_j - p))dt,\ j = 1, ..., m.$$

Using again the Mean Value Theorem, (3.2) and Lemma 2.2, we get

$$\begin{aligned} \beta\|F'(x)-\Delta_k\| &\le \beta L\left(1+\tfrac{1}{2}\prod_{j=1}^{k-1}\beta\|F'(x)-\Delta_{k-j}\|\right)\|x-p\| \\ &\le \beta Lr\left(1+\tfrac{1}{2}\prod_{j=1}^{k-1}\beta\|F'(x)-\Delta_j\|\right) \end{aligned} \tag{3.3}$$

and

$$x-T(x)=(I-\Delta(x))(x-p),$$

where

$$\Delta(x)=\prod_{k=0}^{m-1}F'(x)^{-1}(F(x)-\Delta_{m-k}).$$

Using the notation $\delta_k(x)=\beta\|F'(x)-\Delta_k\|$ and taking into account that $\|F'(x)-\Delta_1\|\le\beta Lr$, we can write

$$\begin{aligned} \delta_1(x) &\le \beta Lr; \\ \delta_k(x) &\le \beta Lr\left(1+\tfrac{1}{2}\prod_{j=1}^{k-1}\delta_j(x)\right),\ k=2,3,..., \end{aligned}$$

and $\|\Delta(x)\|=\prod_{k=1}^{m}\delta_k(x)$. Let $\{f_k\}$ be the numerical sequence obtained from (1.2) for $y=\beta Lr$. It can be easily proved that $\delta_k(x)\le f_k,\ k=1,3,...,m$. Therefore

$$\|\Delta(x)\|=\prod_{k=1}^{m}\delta_k(x)\le\prod_{k=1}^{m}f_k=g_m(\beta Lr).$$

As $\beta Lr<\alpha$, from Lemma 2.1 (b) we have that $g_m(\beta Lr)<\alpha$ and

$$\|\Delta(x)\|<\alpha,\ \forall x\in B(p,r).$$

Now, from $\|\Delta(x)\|\le\alpha$, using the Banach lemma, we have that $I-\Delta(x)$ is invertible and $\|(I-\Delta(x))^{-1}\|\le 1/(1-\alpha)$. Thus, since $x-T(x)=(I-\Delta(x)(x-p)$, we obtain

$$\|x-p\|\le\|(I-\Delta(x))^{-1}\|\|x-T(x)\|\le\frac{1}{1-\alpha}\|x-T(x)\|,\ \forall x\in B(p,r).$$

Therefore p is the unique fixed point of T in $B(p,r)$.

The convergence of the sequence generated by $x_{n+1}=T(x_n)$ results from

$$\|T(x)-p\|=\|\Delta(x)(x-p)\|\le\alpha\|x-p\|.$$

In order to obtain the rate of convergence, from (3.3) and Lemma 2.1 (a), we have

$$\beta\|F'(x)-\Delta_k\|\le\beta L\left(1+\frac{1}{2}\prod_{j=1}^{k-1}f_j(\alpha)\right)<\beta L\left(1+\frac{1}{2}f_1\right)=\frac{\beta L}{\alpha}.$$

Thus $\Delta(x)\| \leq (\beta L/\alpha)^m \|x-p\|^m$ and

$$\|T(x)-p\| \leq \left(\frac{\beta L}{\alpha}\right)^m \|x-p\|^{m+1},\ \forall x \in B(p,r).$$

□

The condition $\|F'(x)^{-1}\| \leq \beta,\ \forall x \in B(p,r)$, and the use of β in the formula of the radius estimation seems to be unusual. More frequently used is the condition $\|F'(p)^{-1}\| \leq \beta_0$ (for instance, the two formulas presented in the Introduction use this condition). This shortcoming can be avoided by replacing the formula $r \leq \alpha/(\beta L)$ with

$$r_0 \leq \frac{\alpha}{(1+\alpha)\beta_0 L} \tag{3.4}$$

where α is the unique positive solution of the equation $x^3 + 2x^2 - 2 = 0$ ($\alpha = \sqrt{3} - 1$). Indeed, take $\beta = \alpha/[(1+\alpha)\beta_0 L]$ and then, as $\beta_0 L r_0 < 1$, using the Banach lemma, we have $\|F'(x)^{-1}\| \leq \beta$. It is worth noticing that this formula is identical with the formula of Hernandez-Romero for Potra-Ptak method, except the equation giving ζ.

We have in turn

Corollary 3.2. *Let p be a solution of $F(x) = 0$ and suppose that $B(p, r_m) \subset C$, where r_m is defined by (3.1) for $\widetilde{\alpha} = \alpha$ and $\widetilde{r} = r$. Suppose that there exists $F'(p)^{-1}$, $\|F'(p)^{-1}\| \leq \beta_0$ and that F' is L-Lipschitz continuous on C. Let r_0 be given by (3.4). Then the sequence $\{x_n\}$ given by the Modified Newton method with starting point $x_0 \in B(p, r_0)$, remains in $B(p, r_0)$ and converges to the unique solution p in $B(p, r_0)$. The rate of convergence is at least $m+1$.*

In the particular case of Potra-Ptak method the formula (3.4) is similar with the formula proposed by Hernandez-Romero $r \leq \frac{\zeta_0}{(1+\zeta_0)\beta_0 L}$, where ζ_0 is the unique positive solution of the equation $x^3 + 4x^2 - 8 = 0$.

It is worth noticing that the estimations proposed by Hernandez-Romero, Catinas and ours (Theorem 3.1), appear to be not comparable. Several numerical experiments show that the values given by these formulas can not be correlated to each other. Table 1 provides the results obtained by applying the Potra-Ptak method for the functions: $f_1(x) = 0.2x^5 - 2x^2 + x,\ f_2(x) = 0.2x^3 - 0.3x^2 + x,\ f_3(x) = x - cos(x)$.

The sign $\star$ means that the derivative of T (the iteration function in Catinas formula) is not Lipschitz continuous around $p = 0$. It can be seen the absence of any order in size between the estimation given by the three formulas.

	Hernandez-Romero	Catinas	Our proposal	The best radius
f_1	0.139634	$\star$	0.109787	0.144933
f_2	0.473295	1.791662	0.512257	∞
f_3	0.925150	0.594361	0.735024	$\approx$1.4806

Table 1: Local radii of cconvergence

Remark 3.1. The radius of convergence proposed in [4] ($r \leq \alpha_m/(\beta L)$, where α_m is the smallest positive solution of the polynomial $g_m(y) - \eta$), depends to some extent on the parameter η. The numerical experiments show that this dependence is not very strong. More precisely, the solution of the equation $g_m(y) - \eta = 0$ is not very sensitive to η for large values of the parameter m. The sequence $\{\alpha_m\}$ is increasing for $\eta < \sqrt{3} - 1$ and decreasing for $\eta > \sqrt{3} - 1$ and tends to $\sqrt{3} - 1$. For example, for $\eta = 0.5$, $\alpha_{10} = 0.73381207...$, $\alpha_{20} = 0.73219954..., \alpha_{30} = 0.73206452..., \alpha_{40} = 0.73205208... \approx \sqrt{3} - 1$.

In what follows we estimate the efficiency of the Modified Newton method and compare it with the efficiency of the Newton method. We will use the classical Ostrowski's index of efficiency, defined by $Index = \varrho^{1/d}$, where ϱ is the convergence order and d is the number of functional evaluations. The computation of $F(x)$ for a given x means n functional evaluations, therefore the computation of the Jacobian of F means n^2 evaluations. Let g and f be the efficiency indexes of Modified Newton method and of Newton method, respectively. We have

$$g(n) = (m+1)^{\frac{1}{n^2+mn}}, \quad f(n) = 2^{\frac{1}{n^2+n}}.$$

For $m = 4$ the graphs of the functions g and f are given in Figure 1.

It can be seen that the efficiency index of Modified Newton method is greater than that of Newton method if n is grater than the abscissa x_P of the point P (in this case, $x_P = 1.269...$). The formula for x_P is

$$x_P(m) = \frac{ln\frac{2^m}{m+1}}{ln\frac{m+1}{2}}.$$

When m increases, the point P moves to the right, but not very fast. For example, if $m = 2$ then $x_P = 0.709...$, if $m = 5$ then $x_P = 1.523...$, if $m = 10$ then $x_P = 2.659...$, etc.

Remark 3.2. The present investigation gives a satisfactory answer to the following problem:

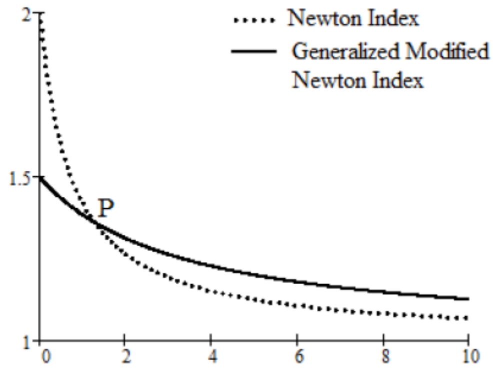

Figure 1: The graphs of efficiency indexes for Modified Newton method (m=4) and for Newton method.

For a given nonlinear equation, if the number of steps m, after which the derivative is re-evaluated, is very large, then the efficiency index of Modified Newton method is decreased. The problem is: *Does there exist a threshold for this number?* The answer is affirmative. If n denotes the unknowns number, such a threshold is the solution m_s of the equation

$$x_P(m) - n = 0.$$

For example, if $n = 3$ the threshold is $m_s = 11$. This means that for equations with $n \geq 3$ the number of steps after which the Jacobian is re-evaluated should be smaller than 11 to obtain an improved efficiency index.

References

[1] **E. Catinas**, Estimating the radius of an attraction ball, *Applied Mathematics Letters*, **22**, (2009), 712-714

[2] **J. A. Ezquerro and M. A. Hernández**, An optimization of Chebyshev's method, *Journal of Complexity*, **25**, (2009), 343-361

[3] **M. A. Hernández-Veron and N. Romero**, On the Local Convergence of a Third Order Family of Iterative Processes, *Algorithms*, **8**, (2015), 1121-1128

[4] **St. Maruster**, Estimating local radius of convergence, *Symposium "Symbolic and Numeric Algorithm for Scientific Computation" (SYNASC), Workshop Iteratime Approximation of Fixed Points, 24-27 Sept. (2016) West University of Timisoara, Timisoara, Romania*

[5] **J. M. Ortega and W. C. Rheinboldt**, *Iterative solution of nonlinear equation in several variables*, Acad. Press, New York, 1970

[6] **F. A. Potra and V. Ptak**, Pitman, London, 1984

[7] **F. Soleymani**, Optimal eighth-order simple root-finders free from derivative, *WSEAS Transactions on Information Science and Applications*, **8 No. 8**, (2011), 293-299

[8] **R. Thukral**, New modification of Newton with third order of convergence for solving nonlinear equation of type $f(0) = 0$, *Amer. J. Comput. Appl. Math.*, **6 No.1**, (2016), 14-18

[9] **J. F. Traub**, *Iterative methods for the solution of equations*, Chelsea Publishing Company, New York, 1982

Ştefan Măruşter

Department of Computer Science
West University of Timisoara
B-l V. Parvan nr. 4
Timisoara
Romania

E-mail: `stefan.maruster@e-uvt.ro`