CONVERGENCE ACCELERATION OF A GENERAL NEWTON METHOD FOR SYSTEMS OF NONLINEAR EQUATIONS

Zorana Lužanin and Sanja Rapajić

Received December 22, 2000

ABSTRACT. The general Newton method proposed in [13] is q-linear convergent. We define a modification of the general Newton method. The main idea of this method is to use a given number of inner iterations for calculating approximation of the inverse Jacobian matrix, because it has an important influence on the convergence rate of the method. Local q-linear, q-superlinear and q-quadratic convergence of the modification of the general Newton method is proved. Some numerical experiments are also presented.

1 Introduction Consider the system of nonlinear equations

where $F : \mathcal{R}^n \to \mathcal{R}^n$ is a nonlinear mapping with the basic assumptions: there exists an $x^* \in \mathcal{R}^n$ with $F(x^*) = 0$, F is continuously differentiable in a neighbourhood of x^* , F' is Lipschitz continuous at x^* and $F'(x^*)$ is nonsingular.

The best known method for finding solution to (1) is the Newton method defined by

Algorithm 1: Newton method (N)

Let $x^0 \in \mathbb{R}^n$ be given.

For $k = 0, 1, 2, \dots$

Step 1: compute s^k as solution of linear system: $F'(x^k)s^k = -F(x^k)$

Step 2: define a new approximation: $x^{k+1} = x^k + s^k$.

The Newton method is attractive because it converges rapidly for any sufficiently good initial guess. In spite of its q-quadratic convergence, this method has a number of disadvantages in practice. Computing all elements of the Jacobian matrix and finding the exact solution of linear system using new matrix for every iterate can be expensive, if the number of unknowns is large.

The first modification of the Newton method is the quasi-Newton method. Basic idea of this method is to eliminate computation of the Jacobian matrix in every iterate. These methods substitute Jacobian matrix $F'(x^k)$ with matrix B_k which is simpler to compute while at the same time gives a relatively good approximation of Jacobian. Instead of matrix B_k , matrix H_k ($H_k = B_k^{-1}$) which approximates matrix $F'(x^k)^{-1}$, can be considered. During the last three decades a large number of methods from this class have been developed. One of the first results on this subject was published by Broyden in 1965, [1, 2]. Substantial contribution to the development of these methods can also be attributed to Dennis, Moré, Martínez and some other [4, 5, 11, 9].

²⁰⁰⁰ Mathematics Subject Classification. 65H10.

 $Key \ words \ and \ phrases.$ nonlinear system, Newton method, general Newton method, rate of convergence.

Algorithm 2: Inverse Broyden method (IB) [1965]

Let $x^0 \in \mathbb{R}^n, H_0 \in \mathbb{R}^{n \times n}$ be given. For k = 0, 1, 2, ...Step 1: define new approximation : $x^{k+1} = x^k - H_k F(x^k)$

Step 2: define a new approximation of Jacobian:

$$H_{k+1} = H_k + \frac{(s^k - H_k y^k)(s^k)^T H_k}{(s^k, H_k y^k)},$$
 where $y^k = F(x^{k+1}) - F(x^k)$ and $s^k = x^{k+1} - x^k$.

In recent years, powerful software packages for automatic differentiations have been developed, so that computing Jacobian matrix does not seem to be a serious problem in most cases. The most important drawback of the Newton method is still the necessity to solve the system of linear equations $(F'(x^k)s^k = -F(x^k))$. Therefore it seems reasonable to use an iterative method to solve the linear system approximately, which represents the inexact Newton method. This method was defined by Dembo, Eisenstat and Steihaug [3].

Algorithm 3: Inexact-Newton method (IN) [1983]

Let $x^0 \in \mathbb{R}^n$ and sequence of real numbers $\{\eta_k\}, \eta_k \geq 0$ be given.

For $k = 0, 1, 2, \dots$

Step 1: find s^k that satisfies:

$$||F'(x^{k})s^{k} + F(x^{k})|| \le \eta_{k}||F(x^{k})||$$

Step 2: define a new approximation: $x^{k+1} = x^k + s^k$.

2 The General Newton method and its modification The basic idea of the general Newton method [13] is to replace the inverse Jacobian matrix $[F'(x^k)]^{-1}$ by its approximation $H(x^k)$, where $H(x^k)$ is invertible. This method, like the inexact Newton methods and the quasi-Newton methods consists of inner and outer iterative methods. The outer iterative method calculates approximation x^k of the exact solution of the system (1), while the inner iterative method computes approximation H_k of the inverse Jacobian matrix $F'(x^k)^{-1}$ using the following iterative method

(2)
$$X^{(p+1)} = X^{(p)} [2I - F'(x)X^{(p)}], \quad p = 0, 1, 2...$$

The important question is the choice of the initial matrix $X^{(0)}$ and the stopping criterion. The general Newton method is defined in [13] by the following algorithm

Algorithm 4: General Newton method (GN) [1996]

Let $x^0 \in \mathbb{R}^n, \ C \in \mathbb{R}^{n \times n}$ such that (3) $|c_{ij}| < \frac{1}{n}, \ i = 1, ..., n, \ j = 1, ..., n.$

and $\varepsilon > 0$ be given.

For $k = 0, 1, 2, \dots$

Step 1: choose matrix $X^{(0)}$ satisfying

$$C = I - F'(x^k)X^{(0)},$$

and calculate $X^{(p+1)}$ by (2). When the stopping criterion

(4)
$$\max |x_{ij}^{(p+1)} - x_{ij}^{(p)}| < \varepsilon$$

is satisfied, put $H_k = X^{(p+1)}$.

Step 2: define a new approximation: $x^{k+1} = x^k - H_k F(x^k)$.

Local q-linear convergence in infinity norm of the general Newton method is proved in [13].

Theorem 1 [13] Suppose that F is two times continuously differentiable on $U(x^*) = \{x, \|x - x^*\|_{\infty} \leq \delta\} \subset D$. Then for arbitrary $x^0 \in U(x^*)$ the sequence $\{x^k\}$ generated by Algorithm 4 remains in $U(x^*)$ and an estimate of the form

$$||x^{k+1} - x^*||_{\infty} \le L ||x^k - x^*||_{\infty}, \quad k = 0, 1, \dots$$

holds for $x^k \in U(x^*)$ and a constant L with $0 \leq L < 1$, provided that δ is sufficiently small.

The main idea of this paper is acceleration of the general Newton method. In order to do that we modify the Algorithm 4. Instead of the stopping criterion (4) the given number of inner iterations are used for calculation of the approximation of the inverse Jacobian matrix.

Now, we define a modification of the general Newton method.

Algorithm 5: Modification of the general Newton method (MGN)

Let $x^0 \in \mathbb{R}^n$ and sequence of nonnegative integers $\{n_k\}, n_k \in \mathbb{N}_0$ be given.

For $k = 0, 1, 2, \dots$

Step 1: choose matrix $X^{(0)}$ such that

$$||C|| = ||I - F'(x^k)X^{(0)}|| < 1,$$

and calculate $X^{(n_k)}$ by (2). Then put $H_k = X^{(n_k)}$

Step 2: define new approximation: $x^{k+1} = x^k - H_k F(x^k)$.

3 Local convergence We need some theorems about the inexact Newton method. Throughout this section $\|\cdot\|$ is used to denote any norm.

The following theorem is proved in [8].

Theorem 2 [8, pp. 99] Let the basic assumptions hold. Then there is ε such that if $||x^0 - x^*|| \le \varepsilon$ and $\{\eta_k\} \subset [0, \eta]$ with $\eta < \overline{\eta} < 1$ then the sequence of inexact Newton iterates $\{x^k\}$ (defined by Algorithm 3) converges q-linearly to x^* with respect to $|| \cdot ||_*$, where $|| \cdot ||_* = ||F'(x^*) \cdot ||$. Moreover,

- if $\lim_{k\to\infty} \eta_k = 0$ the convergence is q-superlinear, and

- if $\eta_k \leq K_{\eta} \|F(x_k)\|^{\alpha}$ for some $K_{\eta} > 0$ then the convergence is q-superlinear with q-order $1 + \alpha$.

We prove the local q-linear convergence of the general Newton method and its modification in the following theorem.

Theorem 3 Assume that the basic assumptions are satisfied and ||C|| < t < 1. Then there exists δ such that if $||x^0 - x^*|| \leq \delta$ then the sequence $\{x^k\}$ obtained by Algorithm 4 and Algorithm 5 converges q-linearly to x^* with respect to $|| \cdot ||_*$.

Proof. First, we prove the q-linear convergence of the method obtained by Algorithm 5.

$$\begin{aligned} \|F'(x^k)s^k + F(x^k)\| &= \|-F'(x^k)H(x^k)F(x^k) + F(x^k)\| \\ &\leq \|I - F'(x^k)H(x^k)\|\|F(x^k)\| \\ &= \|I - F'(x^k)X^{(n_k)}\|\|F(x^k)\|. \end{aligned}$$

It can be proved by mathematical induction that

$$X^{(n_k)} = [F'(x^k)]^{-1}(I - C^{2^{n_k}}).$$

From this we obtain

(5)
$$\begin{aligned} \|F'(x^{k})s^{k} + F(x^{k})\| &\leq \|I - F'(x^{k})[F'(x^{k})]^{-1}(I - C^{2^{n_{k}}})\|\|F(x^{k})\| \\ &= \|C^{2^{n_{k}}}\|\|F(x^{k})\| \\ &\leq \|C\|^{2^{n_{k}}}\|F(x^{k})\| \end{aligned}$$

Since, ||C|| < t < 1 then

$$||C||^{2^{n_k}} < t < 1$$

for every $n_k \in N_0$ and the result follows from the Theorem 2. Analogously, the *q*-linear convergence of the method given by Algorithm 4 can be proved using *p* instead of n_k . \Box

The above theorem supplies q-linear convergence even in the case $H_k = X^{(0)}$ $(n_k = 0)$. From this point of view it is reasonable to expect that inner iterative method for calculating the approximation of the inverse Jacobian has the important influence on the convergence rate of the general Newton method. Namely, suitable choice of the number of inner iterations supplies faster convergence.

Q-superlinear convergence of the method defined by Algorithm 5 follows from the following corollary.

Corollary 1 Assume that the sequence $\{x^k\}$ defined by Algorithm 5 converges to x^* . If

$$\lim_{k \to \infty} n_k = \infty.$$

then the convergence is q-superlinear.

Proof. From the enequality (5) it holds that

$$\|F'(x^k)s^k + F(x^k)\| \le \|C\|^{2^{n_k}} \|F(x^k)\|.$$

Put $\eta_k = \|C\|^{2^{n_k}}$. Since $\|C\| < 1$ it follows that $\lim_{k \to \infty} \|C\|^{2^{n_k}} = 0$, i.e. $\lim_{k \to \infty} \eta_k = 0$ so the result follows from the Theorem 2.

From the Corollary 1 it is easy to determine some sequences $\{n_k\}$ for which the method defined by Algorithm 5 converges q-superlinearly. We test these choices of n_k :

(6)
$$n_k = k + 1, \qquad n_k = \lfloor \sqrt{k} \rfloor + 1.$$

Now, let

(7)
$$n_k = \lfloor \frac{\ln \|F(x^k)\|}{\ln \|C\|} \rfloor.$$

Corollary 2 Assume that the sequence $\{x^k\}$ defined by Algorithm 5 with $\{n_k\}$ given by (7) converges to x^* . Then the convergence is q-quadratic.

Proof. Let $\eta_k = ||C||^{2^{n_k}}$. Then we have

$$\eta_k < \|C\|^{n_k} = \|C\|^{\lfloor \frac{\ln \|F(x^k)\|}{\ln \|C\|} \rfloor} = \|C\|^{\lfloor \log_{\|C\|} \|F(x^k)\| \rfloor} \le \|C\|^{\log_{\|C\|} \|F(x^k)\|} = \|F(x^k)\|.$$

Convergence is q-quadratic by Theorem 2.

Numerical results In this section we present some numerical results for the following 4 methods:

N:	Newton method (Algorithm 1);
IB:	Inverse Broyden method (Algorithm 2);
GN:	General Newton method (Algorithm 4);
MGN:	Modification of GN method (Algorithm 5)

The stopping criterion is

$$|F(x^k)|| \le 10^{-10}$$
 and $||x^k - x^{k-1}|| \le 10^{-4} ||x^k|| + 10^{-4}$.

We consider four systems of nonlinear equations. The start approximation and the exact solution are denoted by x^0 and x^* , respectively.

Problem 1: [14]

$$f_1(x) = \sin x_1 + 2x_2 - 1$$

$$f_2(x) = 2x_1 + \cos x_2 - 2$$

 $x^0 = (0,0)$ and $x^0 = (0.5, 0.5)$

Problem 2: [15]

$$f_1(x) = x_1 - 0.7 \sin x_1 - 0.2 \cos x_2$$

$$f_2(x) = x_2 - 0.7 \cos x_1 - 0.2 \sin x_2$$

 $x^0 = (0,0)$ and $x^0 = (0.5, 0.5)$ **Problem 3:** [10]

$$f_1(x) = x_1^3 - 3x_1x_2^2 - 1$$

$$f_2(x) = 3x_1^2x_2 - x_2^3$$

 $x_1^* = (1,0), \, x_2^* = (-0.5, \sqrt{3} \cdot 0.5), \, x_3^* = (-0.5, -\sqrt{3} \cdot 0.5) \\ x^0 = (1.5, 0.5), \, x^0 = (-1,1), \, x^0 = (-2, -1.5) \text{ and } x^0 = (-2, 1.5)$ **Problem 4:** [12] c ()

$$f_1(x) = 2x_1 - x_2 + 0.5h(x_1 + h + 1)^3$$

$$f_i(x) = 2x_i - x_{i-1} - x_{i+1} + 0.5h(x_i + ih + 1)^3 \quad i = 2, 3, \dots, n-1$$

$$f_n(x) = 2x_n - x_{n-1} + 0.5h(x_n + nh + 1)^3$$

where $h=\frac{1}{n+1}~x^0=(0,...,0)$ and $x^0=(0.5,...,0.5)$

For methods defined by Algorithm 4 (GN) and Algorithm 5 (MGN) the constant matrix $C = \begin{bmatrix} 0.2 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$ given in [13] is used in problems 1-3. We also use $\varepsilon = 0.1$ for the general Newton method. The following table presents the number of outer iterations needed for convergence. It can be seen that the convergence order of the MGN method depends on the choice of n_k .

				$\mathrm{MGN}~(n_k)$				
problem	x^0	Ν	IB	1	k + 1	$\lfloor k \rfloor + 1$	$\left\lfloor \frac{\ln \ F(x^k)\ }{\ln \ C\ } \right\rfloor$	GN
1	(0,0)	4	6	10	4	5	4	10
1	(0.5, 0.5)	4	6	9	4	5	4	9
2	(0,0)	5	8	10	4	5	4	5
2	(0.5, 0.5)	4	6	9	4	5	3	4
3	(1.5, 0.5)	6	11	10	6	6	6	10
3	(-1,1)	5	10	11	5	6	5	11
3	(-2, -1.5)	7	15	13	7	7	7	13
3	(-2, 1.5)	7	15	12	7	7	7	12

Table 1.

In Table 2. we present some iteration statistics for problem 3 with initial guess $x^0 = (1.5, 0.5)$. We tabulate the iteration counter and the ratio $\frac{||x^{k+1}-x^*||_2}{||x^k-x^*||_2}$ for the GN and MGN methods in this table.

k	1	k + 1	$\lfloor k \rfloor + 1$	$\left\lfloor \frac{\ln \ F(x^k)\ }{\ln \ C\ } \right\rfloor$	GN
0	0.421203	0.421203	0.421203	0.421203	0.421203
1	0.257178	0.240545	0.240545	0.257178	0.257178
2	0.0573532	0.0730239	0.0694743	0.0573532	0.0573532
3	0.118087	0.00526551	0.00246931	0.00440029	0.118087
4	0.681367	$2.75 10^{-5}$	$5.83 \ 10^{-5}$	$1.93 \ 10^{-5}$	0.681367
5	0.0895832	$2.26 \ 10^{-11}$	$6.48 \ 10^{-5}$	$3.45 \ 10^{-10}$	0.0895832
6	0.089995	$1.76 \ 10^{-16}$	$4.64 \ 10^{-5}$	$1.87 \ 10^{-16}$	0.089995
	linear	superlin.	superlin.	superlin.	linear
		or better	or better	or better	

Table 2.

We consider the problem 4 in Table 3. Dimension of the system is denoted by n. For this problem the initial approximation $X^{(0)}$ of the inverse Jacobian matrix in k + 1-th iteration is determined by matrix H_k from k-th iteration.

					MGN (n_k)			
x^0	n	Ν	IB	1	k + 1	$\lfloor k \rfloor + 1$	$\lfloor \frac{\ln \ F(x^k)\ }{\ln \ C\ } \rfloor$	GN
(0,0)	8	5	10	6	5	5	5	5
(0,0)	32	6	15	7	6	6	6	6
(0.5, 0.5)	8	5	13	7	5	5	5	6
(0.5, 0.5)	$\overline{32}$	6	19	8	6	7	6	6

Table 3.

The results given in Table 3. show that the choice of $X^{(0)}$ can also accelerate the convergence of the GN and MGN methods. In this example the GN and MGN methods perform similarly to the Newton method. This can be interesting for further research.

References

- G. Broyden, A class of methods for solving nonlinear simultaneous equation, Math. Comp. 19(1965), 577-593.
- [2] G. Broyden, The convergence of single-rank quasi-Newton methods, Math. Comp. 24(1970), 365-382.
- [3] R. S. Dembo, S. C. Eisenstat, T. Steihaug, Inexact Newton methods. SIAM J. Numer. Anal. 19(1982), 400-408.
- [4] J. E. Dennis, J. J. Moré, Quasi-Newton method, motivation and theory, SIAM Review, 19(1977), 46-89.
- [5] J. E. Dennis, G. Li, Hybrid algorithm for solving sparse nonlinear systems of equations, Math. Comp., 50(1988), 155-166.
- [6] J. E. Dennis, R. B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall, Englewood Cliffs, NJ, 1983.
- [7] S. C. Eisenstat, H. F. Walker, Choosing the forcing terms in an inexact Newton method, SIAM J.Sci.Comput., 17(1996), 16-32.
- [8] C. T. Kelley, Iterative Methods for Linear and Nonlinear Equations, Society for Industrial and Applied Mathematics, Philadelphia, 1995.
- [9] N. Krejić, Z. Lužanin, Newton-like method with modification of the right-hand side vector, Math. Comp., (accepted for publication).
- [10] Z. Lužanin, Global and Local Convergence of Modifications of Newton Method, Ph. D. Thesis, University of Novi Sad, 1997.
- [11] J. M. Martínez, A family of quasi-Newton methods for nonlinear equations with direct secant updates of matrix factorizations, SIAM J.Numer.Anal. 27(1990), 1034-1049.
- [12] J. J. Moré, M. Y. Cosnard, Numerical solution of nonlinear equations, AMC Trans. Math. Software 5(1979), 64-85.
- [13] T. Noda, T. Fujita, A general Newton method for systems of nonlinear equations, Math. Japonica 44, No. 3(1996), 541-547.
- [14] D. L. Russell, Optimization Theory, Benjamin, New York, 1970.
- [15] H. Sunouchi, Numerical Calculus, (in Japanese), SAIENSU-SHA, 1987.

Institute of Mathematics, University Novi Sad, Trg Dositeja Obradovića 4, 21000 Novi Sad, Yugoslavia

e-mail address: zorana@unsim.im.ns.ac.yu, sanja@unsim.im.ns.ac.yu