Journal of Applied
and Numerical Analysis

2023, Vol. 1, pp. 80–87
doi: 10.30970/ana.2023.1.80

On the Newton-Broyden method for solving systems of nonlinear equations

S. Shakhno, H. Yarmola

Ivan Franko National University of Lviv,
1, Universytets'ka str., 79000, Lviv, Ukraine

Анотація. У роботі досліджено метод Ньютона-Бройдена для розв'язування систем нелінійних рівнянь з декомпозицією оператора. Проведено аналіз локальної збіжності за класичних умов Ліпшиця та наведено результати чисельних експериментів. Продемонстровано ефективність цього методу для чисельного розв'язування різних типів нелінійних систем.

Abstract. The Newton-Broyden method for solving systems of nonlinear equations with operator decomposition is studied in this paper. A local convergence analysis is provided under classical Lipschitz conditions and results of numerical experiments are presented. The effectiveness of this method in solving various types of nonlinear systems is demonstrated.

1 Introduction

Consider a system of nonlinear equations

$$H(x) = 0, \tag{1.1}$$

where $H : D \to \mathbb{R}^n$ is a nonlinear operator, D is the open convex set of the space $\mathbb{R}^n$. Our goal is to find a solution $x^* \in D$ that satisfies the equation (1.1).

The most used method for solving the system (1.1) is the classical Newton method [5]

$$x_{k+1} = x_k - H'(x_k)^{-1} H(x_k),\ k \ge 0, x_0 \in D,$$

which has a quadratic order of convergence. However, it requires analytically given first-order derivatives. If it is difficult or impossible to do, then we can apply methods with divided differences or the approximation of the matrix of partial derivatives. One such approach is the application of the Broyden's update formula [2, 4, 5]. Broyden method has form

$$\begin{aligned} &x_{k+1} = x_k - B_k^{-1} H(x_k),\ k \ge 0, \\ &B_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^\top}{s_k^\top s_k}, \\ &s_k = x_{k+1} - x_k, \quad y_k = H(x_{k+1}) - H(x_k), \end{aligned} \tag{1.2}$$

where $B_0 \in \mathbb{R}^{n \times n}$ is a given matrix. One iteration of solving the system (1.1) using (1.2) consists of the following steps:

1) solve a linear system

$$B_k s_k = -H(x_k)$$

and calculate x_{k+1} by formula

$$x_{k+1} = x_k + s_k;$$

2) calculate the matrix B_{k+1} according to the Broyden's update formula (1.2).

Key words: Nonlinear system, Broyden method, local convergence, operator decomposition.

Let a nonlinear operator H can be rewritten as the sum of two nonlinear operators F an G and consider the system

$$H(x) \equiv F(x) + G(x) = 0. \tag{1.3}$$

Here F, $G : D \to \mathbb{R}^n$. In [3] there was proposed the Broyden-like method

$$\begin{aligned} x_{k+1} &= x_k - B_k^{-1} H(x_k),\ k \geq 0, \\ B_{k+1} &= B_k + \frac{(y_k - B_k s_k) s_k^\top}{s_k^\top s_k}, \\ s_k &= x_{k+1} - x_k, \quad y_k = F(x_{k+1}) - F(x_k). \end{aligned} \tag{1.4}$$

This method was studied under classical Lipschitz condition in [1, 3]. Moreover, the case of the nondifferentiable operator G was considered.

Suppose that the Jacobian matrix F' is calculated exactly and instead of G' the Broyden's update formula is used. As a result, we obtain the following method, which is built on the basis of both Newton and Broyden methods [8]

$$x_{k+1} = x_k - [F'(x_k) + B_k]^{-1}(F(x_k) + G(x_k)), k = 0, 1, 2, ..., \tag{1.5}$$

where matrices B_k are defined by

$$\begin{aligned} B_{k+1} &= B_k + \frac{(y_k - B_k s_k) s_k^\top}{s_k^\top s_k}, \\ s_k &= x_{k+1} - x_k, \quad y_k = G(x_{k+1}) - G(x_k). \end{aligned} \tag{1.6}$$

In this paper we provide the local convergence analysis of the Newton-Broyden method (1.5)-(1.6) under classical Lipschitz condition. The differentiability of both operators F and G is assumed. Moreover, we show the applicability of the proposed method for solving different types of nonlinear systems. The Newton-Broyden method (1.5)-(1.6) can be applied for solving systems with the nonlinear function containing a nondifferentiable part or if there are difficulties in calculating the derivative. In the second case, the function H can be presented as the sum of two functions F and G, and G contains the problematic part.

2 Convergence Analysis

Let $U(\bar{x}, r) = \{x \in D,\ \|x - \bar{x}\| < r\}$ denotes the open ball with center $\bar{x} \in \mathbb{R}^n$ and the radius r, $\bar{U}(\bar{x}, r)$ – the closed ball, and I denotes the identity matrix. In the theoretical statements l_2-vector and matrix norm is used.

First, we present some lemmas that are needed to obtain the main result.

Lemma 2.1. [3] *Let $s \in \mathbb{R}^n$ and $s^\top s = 1$. Then*

$$\|I - ss^\top\| = 1.$$

Lemma 2.2. *Let $x^* \in D$ be a solution of system (1.3) and $D \subseteq \mathbb{R}^n$ be an open convex domain containing all x_k, and $x_k \neq x^*$. Let $F, G : D \to \mathbb{R}^n$, $F, G \in C^1(D)$, an inverse matrix $H'(x^*)^{-1}$ exists and suppose that for all $x \in D$ the following conditions hold*

$$\|H'(x^*)^{-1}(F'(x) - F'(x^*))\| \leq \gamma \|x - x^*\|, \tag{2.1}$$

$$\|H'(x^*)^{-1}(G'(x) - G'(x^*))\| \leq \eta \|x - x^*\|. \tag{2.2}$$

Let B_{k+1} be defined by the formula (1.6) and $A_{k+1} = F'(x_{k+1}) + B_{k+1}$. Then the following estimates hold for all $k \geq 0$

$$\begin{aligned} \|H'(x^*)^{-1}(A_{k+1} - H'(x^*))\| &\leq \|H'(x^*)^{-1}(B_k - G'(x^*))\| + \gamma \|x_{k+1} - x^*\| \\ &\quad + \frac{\eta}{2}\Big[\|x_{k+1} - x^*\| + \|x_k - x^*\|\Big]. \end{aligned} \tag{2.3}$$

Proof. We can write

$$\begin{aligned}
A_{k+1} - H'(x^*) &= F'(x_{k+1}) - F'(x^*) + B_k - G'(x^*) + \frac{(y_k - B_k s_k)s_k^\top}{s_k^\top s_k} \\
&= F'(x_{k+1}) - F'(x^*) + B_k - G'(x^*) \\
&\quad + \frac{(G'(x^*)s_k - B_k s_k)s_k^\top}{s_k^\top s_k} + \frac{(y_k - G'(x^*)s_k)s_k^\top}{s_k^\top s_k} \\
&= (B_k - G'(x^*))\left[I - \frac{s_k s_k^\top}{s_k^\top s_k}\right] + F'(x_{k+1}) - F'(x^*) \\
&\quad + \frac{(y_k - G'(x^*)s_k)s_k^\top}{s_k^\top s_k}.
\end{aligned}$$

Then

$$\begin{aligned}
\|H'(x^*)^{-1}(A_{k+1} - H'(x^*))\| &\le \|H'(x^*)^{-1}(B_k - G'(x^*))\| \left\|I - \frac{s_k s_k^T}{s_k^\top s_k}\right\| \\
&\quad + \|H'(x^*)^{-1}(F'(x_{k+1}) - F'(x^*))\| \\
&\quad + \left\|\frac{H'(x^*)^{-1}(y_k - G'(x^*)s_k)s_k^\top}{s_k^\top s_k}\right\|.
\end{aligned}$$

Taking into account Lemma 2.1, conditions (2.1), (2.2) and the inequality

$$\begin{aligned}
\|H'(x^*)^{-1}(y_k - G'(x^*)s_k)\| &= \left\| \int_0^1 H'(x^*)^{-1}\Big(G'(x_k + \theta(x_{k+1} - x_k)) - G'(x^*)\Big) d\theta s_k \right\| \\
&\le \frac{\eta}{2}(\|x_{k+1} - x^*\| + \|x_k - x^*\|)\|s_k\|,
\end{aligned}$$

we get the estimate (2.3). □

Now we give the local convergence theorem for the method (1.5)-(1.6).

Theorem 2.3. *Let $F, G : D \to \mathbb{R}^n$ be nonlinear operators, $D \subseteq \mathbb{R}^n$ be the open convex set and $F, G \in C^1(D)$. Suppose that*

1) x^* be a solution of the equation (1.1) and an inverse matrix $H'(x^*)^{-1}$ exists;

2) the Jacobian matrices F' and G' satisfy following Lipschitz conditions for all $x \in D$

$$\|H'(x^*)^{-1}(F'(x) - F'(x^*))\| \le \gamma \|x - x^*\|, \tag{2.4}$$

$$\|H'(x^*)^{-1}(G'(x) - G'(x^*))\| \le \eta \|x - x^*\|. \tag{2.5}$$

Then for all matrix B_0 such that

$$\|H'(x^*)^{-1}(B_0 - G'(x^*))\| \le b = \frac{\gamma + 2\eta}{7\gamma + 13\eta} \tag{2.6}$$

the sequence $\{x_k\}_{k\ge 0}$ generated by the method (1.5)-(1.6) is well-defined, remains in $\bar{U}(x^*, r) \subset D$, where $r \le \dfrac{1}{7\gamma + 13\eta}$, and converges to x^*. Moreover, the following error estimates hold for all $k \ge 0$

$$\|x_{k+1} - x^*\| \le q\,\|x_k - x^*\|, \ q = \frac{1}{2}. \tag{2.7}$$

Proof. We prove this theorem using mathematical induction.

Let $x_0 \in \bar{U}(x^*, r)$. Then, by using conditions (2.4) and (2.6), we have

$$\begin{aligned} \|H'(x^*)^{-1}(A_0 - H'(x^*))\| &\le \|H'(x^*)^{-1}(F'(x_0) - F'(x^*))\| \\ &\quad + \|H'(x^*)^{-1}(B_0 - G'(x^*))\| \\ &\le b + \gamma\|x_0 - x^*\| \le b + \gamma r \le 2b < 1. \end{aligned}$$

It follows by the Banach lemma on the invertible operator (Theorem on the existence and norm of inverse matrix [5]) that A_0^{-1} exists and

$$\|A_0^{-1}H'(x^*)\| \le \frac{1}{1 - b - \gamma\|x_0 - x^*\|} \le \frac{1}{1 - 2b}.$$

So, the iterate x_1 is well-defined and we can write that

$$\begin{aligned} x_1 - x^* &= x_0 - x^* - A_0^{-1}(H(x_0) - H(x^*)) \\ &= A_0^{-1}\{A_0(x_0 - x^*) - (H(x_0) - H(x^*))\} \\ &= A_0^{-1}\Big\{F'(x_0) \mp F'(x^*) - \int_0^1 F'(x^* + \theta(x_0 - x^*))d\theta\Big\}(x_0 - x^*) \\ &\quad + A_0^{-1}\Big\{B_0 \mp G'(x^*) - \int_0^1 G'(x^* + \theta(x_0 - x^*))d\theta\Big\}(x_0 - x^*). \end{aligned}$$

Using (2.4) – (2.6), we obtain

$$\begin{aligned} \|x_1 - x^*\| &\le \frac{1.5\gamma\|x_0 - x^*\|^2 + (b + 0.5\eta\|x_0 - x^*\|)\|x_0 - x^*\|}{1 - (b + \gamma\|x_0 - x^*\|)} \\ &\le \frac{2b + 0.5(\gamma + \eta)\|x_0 - x^*\|}{1 - (b + \gamma\|x_0 - x^*\|)}\|x_0 - x^*\| \le q\|x_0 - x^*\| < r. \end{aligned}$$

Hence, $x_1 \in \bar{U}(x^*, r)$. From the Lemma 2.2 we get

$$B_1 - G'(x^*) = (B_0 - G'(x^*))\left(I - \frac{s_0 s_0^\top}{s_0^\top s_0}\right) + \frac{(y_0 - G'(x^*)s_0)s_0^\top}{s_0^\top s_0}$$

and

$$\begin{aligned} \|H'(x^*)^{-1}(A_1 - H'(x^*))\| &\le b + \gamma\|x_1 - x^*\| + \frac{\eta}{2}\Big(\|x_1 - x^*\| + \|x_0 - x^*\|\Big) \\ &\le b + (\gamma + \eta)r \le 2b. \end{aligned}$$

Suppose that for all $0 \le i \le k - 1$

$$\begin{aligned} \|H'(x^*)^{-1}(A_i - H'(x^*))\| &\le 2b; \\ \|x_{i+1} - x^*\| &\le q\|x_i - x^*\|; \\ x_{i+1} &\in \bar{U}(x^*, r). \end{aligned}$$

Then, for $i = k$ we get

$$\begin{aligned} \|H'(x^*)^{-1}(A_k - H'(x^*))\| &\le \|H'(x^*)^{-1}(B_0 - G'(x^*))\| + \gamma\|x^* - x_k\| \\ &\quad + \eta\sum_{i=0}^{k-1}\|x^* - x_i\| \\ &\le b + \gamma r + \eta r/(1 - q) = b + (\gamma + 2\eta)r \le 2b < 1. \end{aligned}$$

So,

$$\|A_k^{-1}H'(x^*)\| \le \frac{1}{1-b-(\gamma+2\eta)r} \le \frac{1}{1-2b}$$

and

$$\begin{aligned} x_{k+1}-x^* &= x_k - x^* - A_k^{-1}(H(x_k)-H(x^*)) \\ &= A_k^{-1}\{A_k(x_k-x^*) - (H(x_k)-H(x^*))\} \\ &= A_k^{-1}\Big\{F'(x_k) \mp F'(x^*) - \int_0^1 F'(x^*+\theta(x_k-x^*))d\theta\Big\}(x_k-x^*) \\ &\quad + A_k^{-1}\Big\{B_k \mp G'(x^*) - \int_0^1 G'(x^*+\theta(x_k-x^*))d\theta\Big\}(x_k-x^*). \end{aligned}$$

Using (2.4) – (2.6), we obtain

$$\begin{aligned} \|x_{k+1}-x^*\| &\le \frac{2b+0.5(\gamma+\eta)\|x_k-x^*\|}{1-2b}\|x_k-x^*\| \\ &\le q\|x_k-x^*\| < r. \end{aligned}$$

Hence, $x_{k+1} \in \bar{U}(x^*, r)$ and (2.7) hold. Induction is completed.

From the estimate (2.7), taking into account that $q=\frac{1}{2}$, we get that $\lim\limits_{k\to\infty}\|x_k-x^*\|=0$ and the sequence $\{x_k\}_{k\ge 0}$ converges to x^*. □

The estimate (2.7) shows that the sequence $\{x_k\}_{k\ge 0}$ converges to x^* linearly.

3 Numerical Examples

In this section we demonstrate the applicability of the Newton-Broyden method (1.5)-(1.6). We compare this iterative process with the Broyden method (1.2). For this, we present the results obtained for the systems of nonlinear equations with following nonlinear functions [6,7].

Example 3.1. *Trigonometric-exponential function [7].*

$$F(x) = \begin{cases} 3x_i^3 + 2x_{i+1} - 5, & i = 1, \\ 3x_i^3 + 2x_{i+1} + 4x_i - 8, & 1 < i < n, \\ 4x_i - 3, & i = n, \end{cases}$$

$$G(x) = \begin{cases} \sin(x_i - x_{i+1})\sin(x_i + x_{i+1}), & i = 1, \\ \sin(x_i - x_{i+1})\sin(x_i + x_{i+1}) & \\ -x_{i-1}\exp(x_{i-1} - x_i), & 1 < i < n, \\ -x_{i-1}\exp(x_{i-1} - x_i), & i = n. \end{cases}$$

Example 3.2. *Gheri and Mancino problem [6].*

$$F_i(x) = 14nx_i + \Big(i - \frac{n}{2}\Big)^3, \quad 1 \le i \le n,$$

$$G_i(x) = \sum_{j=1, j\ne i}^{n} z_{ij}\Big(\sin^5(\ln(z_{ij})) + \cos^5(\ln(z_{ij}))\Big), \quad 1 \le i \le n,$$

where $z_{ij} = \sqrt{x_j^2 + i/j}$, $1 \le i \le n$, $1 \le j \le n$.

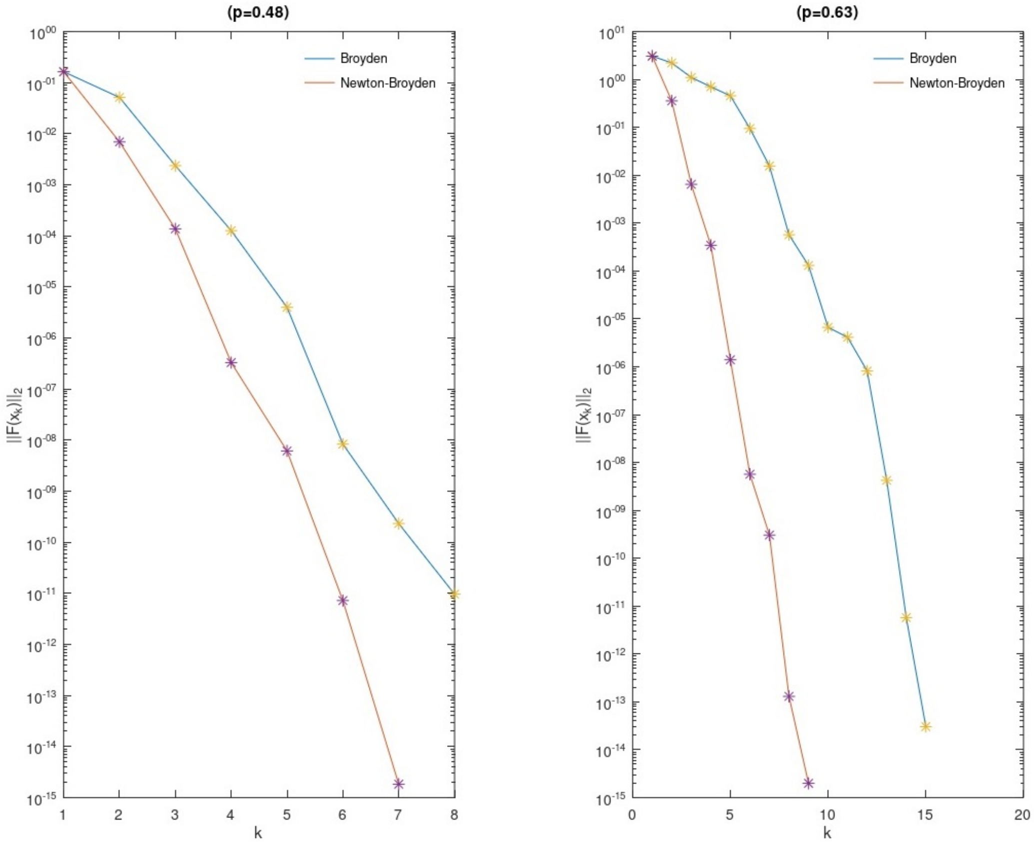

Fig. 3.1. Result for Example 3.3

Example 3.3. *System with non-differentiable functions.*

$$F(x) = \begin{cases} x_3^2(1-x_2) - x_1x_2, \\ x_3^2(x_1^3 - x_1) - x_2^2, \\ x_1 + x_2 + x_3 - 4, \end{cases}$$

$$G(x) = \begin{cases} |x_2 - x_3^2|, \\ |6x_2 - x_3^2 - x_1|, \\ \ln|x_1|. \end{cases}$$

To stop iterative processes we use condition

$$\|x_{k+1} - x_k\| \le 10^{-10} \quad \text{and} \quad \|H(x_{k+1})\| \le 10^{-10}.$$

The matrix B_0 was calculated as a matrix of first-order divided difference at the points x_0 and $y_0 = x_0 + 10^{-4}$.

Table 3.1. The number of iterations for Example 3.1

Method	$p = 0.6$	$p = 1$	$p = 2$
Broyden	11	24	59
Newton-Broyden	7	13	17

Tables 3.1 and 3.2 show the results obtained for Example 3.1 and 3.2 with 50 equations, respectively. The initial approximation for Example 3.1 is $x_0 = (2, \ldots, 2)^\top p$, where p is a real number,

and the exact solution is $x^* = (1, \dots, 1)^\top$. The initial approximate for Example 3.2 is $x_0 = (1, \dots, 1)^\top p$, where p is a real number.

Table 3.2. The number of iterations for Example 3.2

Method	$p = 0$	$p = 10$	$p = 20$
Broyden	7	7	8
Newton-Broyden	7	7	8

Table 3.3. The number of iterations for Example 3.3

Method	$p = 0.48$	$p = 0.63$	$p = 0.4$
Broyden	8	15	13
Newton-Broyden	7	9	11

In Table 3.3 and Figure 3.1, there are results for Example 3.3. The initial approximate is $x_0 = (-2,\ 4,\ 6)^\top p$ and the exact solution is $x^* = (-1,\ 2,\ 3)^\top$. Figure 3.1 demonstrates the change of the residual's norm at each step.

From the obtained results, we can see that the Newton-Broyden method (1.5)-(1.6) has an advantage in the number of iterations over the Broyden method. For Example 3.2, the number of iterations is the same for both methods. This is explained by the linearity of the operator F.

4 Conclusions

In this article we studied the local convergence of the Newton-Broyden method for solving systems of nonlinear equations. Convergence analysis was provided under the classical Lipschitz conditions for the first-order derivatives of the nonlinear operators. The applicability of this method is confirmed by numerical examples. The Newton-Broyden method can be applied for systems with nondifferentiable functions.

References

1. Argyros, I.K.: On the convergence of Broyden-Like methods. Acta Mathematica Sinica, English Series. **23** (11), 2087-2096 (2007). doi: 10.1007/s10114-007-0963-3
2. Broyden, C.G.: A class of methods for solving nonlinear simultaneous equations. Mathematical Computations. **19**, 577-593 (1965). doi: 10.1090/S0025-5718-1965-0198670-6
3. Chen, X.: On the convergence of Broyden-like methods for nonlinear equations with non-differentiable terms. Annals of the Institute of Statistical Mathematics. **42**, 387-401 (1990). doi: 10.1007/BF00050844
4. Dennis, Jr., J.E.: On the convergence of Broyden's method for nonlinear systems of equations. Mathematics of Computation. **25** (115), 559-567 (1971). doi: 10.2307/2005218
5. Dennis, Jr., J.E., Schnabel, R.B.: Numerical methods for unconstrained optimization and non-linear equations. Prentice-Hall, Englewood Cliffs (1983)
6. Gheri, G., Mancino, O.G.: A significant example to test methods for solving systems of nonlinear equations. Calcolo. **8**, 107-113 (1971). doi: 10.1007/BF02575578
7. Lukšan, L.: Inexact trust region method for large sparse systems of nonlinear equations. Journal of Optimization Theory and Applications. **81**, 569–590 (1994)

8. Shakhno, S.M., Bodnar, O.Ya., Yarmola, H.P.: Combined Newton-Broyden method for solving nonlinear equations. In: Abstract of conference “Actual problems of applied mathematics and computer science”, Lviv, 333-334 (2015)

Received 29.10.2023
Revised 09.11.2023