NEWTON-LIKE METHOD WITH MODIFICATION
OF THE RIGHT-HAND-SIDE VECTOR

NATAŠA KREJIĆ AND ZORANA LUŽANIN

ABSTRACT. This paper proposes a new Newton-like method which defines new iterates using a linear system with the same coefficient matrix in each iterate, while the correction is performed on the right-hand-side vector of the Newton system. In this way a method is obtained which is less costly than the Newton method and faster than the fixed Newton method. Local convergence is proved for nonsingular systems. The influence of the relaxation parameter is analyzed and explicit formulae for the selection of an optimal parameter are presented. Relevant numerical examples are used to demonstrate the advantages of the proposed method.

1. INTRODUCTION

Consider the system of nonlinear equations

$$(1.1) \qquad F(x) = 0$$

where $F : R^n \to R^n$ is a nonlinear C^1-function.

The best-known method for finding a solution to (1.1) is the Newton method.

Algorithm N: (Newton method)

Let x^0 be given.

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

- Step 1. Solve

$$(1.2) \qquad F'(x^k)s_N^k = -F(x^k).$$

- Step 2. Define $\quad x^{k+1} = x^k + s_N^k.$

Equation (1.2) is known as the Newton system, while the vector s_N^k is called the Newton correction. The Newton method possesses good theoretical characteristics such as quadratic convergence for any sufficiently good initial guess. Despite its obvious qualities, this method has a number of disadvantages in practice. This is primarily the case with the high cost of a Newton iterate, due to the necessity to compute all the elements of the Jacobian matrix, as well as the need for an exact solution of a system of linear equations using a new matrix for every iterate.

In recent years, powerful software packages for automatic differentiation have been developed (see [7] and the references therein), so that computing analytic derivatives in an efficient way in many cases does not seem to present serious problem. In these cases the most important drawback of the Newton method is the

necessity to solve the Newton system. A possible way of overcoming this difficulty is to apply the inexact Newton method, [4, 8], and to solve the Newton system only approximately by means of some iterative method.

$\quad$ *Algorithm IN:* (Inexact Newton method)

Let x^0 be given.

$\quad$ For $k = 0, 1, 2, \ldots$

$\bullet$ Step 1. Find some step s^k which satisfies

$$(1.3) \qquad\qquad F'(x^k)s^k = -F(x^k) + r^k$$

$\quad$ where

$$||r^k|| \leq \eta_k ||F\left(x^k\right)||.$$

$\bullet$ Step 2. Define $\quad x^{k+1} = x^k + s^k$.

Another possibility is to use some quasi-Newton method "with cheap linear algebra". The simplest modification of the Newton method is the fixed Newton method (also known as the chord method or modified Newton method).

$\quad$ *Algorithm FN:* (Fixed Newton method)

Let x^0 be given.

$\quad$ For $k = 0, 1, 2, \ldots$

$\bullet$ Step 1. Solve

$$(1.4) \qquad\qquad F'(x^0)s_F^k = -F(x^k).$$

$\bullet$ Step 2. Define $\quad x^{k+1} = x^k + s_F^k$.

This method diminishes both disadvantages of the Newton method mentioned above but is significantly slower. If LU decomposition is used as a linear solver, then solving of (1.4) requires $\mathcal{O}\left(n^2\right)$ flops (except for the first iteration), as opposed to (1.2) where $\mathcal{O}\left(n^3\right)$ flops are needed. From the computational complexity point of view, the fixed Newton method is significantly cheaper. However, it is only linearly convergent. That is why we can say that the Newton method is "expensive and fast" while the fixed Newton method is "cheap and slow".

In this paper we propose a new method, the method of Modification of the Right-hand-side Vector (MRV method), which is cheaper than the Newton and faster than the fixed Newton method. The main idea is to modify the right-hand-side vector in (1.4) so that the fixed Newton method becomes more similar to the Newton method. Contrary to the common practise in Newton-like methods, we do not update the approximation of the Jacobian in every iteration, but keep the initial approximation through the whole iteration process fixed and update the right-hand-side vector in each iteration using the Jacobian. In this way, one step of the method requires the solution of a linear system, as is usual with this type of methods, but the system matrix is the same in all iterations. In fact, the inverse Jacobian is updated in every iteration with the perturbation of the right-hand-side vector as will be explained later. The first attempt of this kind was made in [12], but with much less success than here.

This paper is organized as follows. In Section 2 we define the MRV method and analyze its local convergence for the system (1.1) assuming a suitable choice of the initial approximation and the relaxation parameter. The choice of the relaxation parameter is analyzed in Section 3. The optimal choice is the solution of the minimization problem. In Section 4 we describe our test functions and present a

numerical comparison of our new method with some of the well-known methods. Finally, in Section 5 we conclude with some final remarks and suggestions for future research.

2. Convergence result for the new method

In this section we will consider the method for solving the system (1.1) where F is a continuously differentiable mapping. The method is given by the following algorithm.

    Algorithm MRV:

Let $x^0 \in R^n$ and $A \in R^{n \times n}$ nonsingular matrix be given.

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

- Step 1. Solve

$$(2.1) \qquad As^k = (\alpha_k H(x^k) - E)F(x^k).$$

- Step 2. Define $x^{k+1} = x^k + s^k$,

where

$$H(x^k) = F'(x^k) - A,$$

E is the identity matrix and α_k is a real parameter. In every iteration one has to calculate $H(x^k)$, assuming that the Jacobian is available, and to solve the system of linear equations (2.1). Generally, in this way the cost of one iteration, in general, is $\mathcal{O}\left(n^2\right)$, since the LU factorization of A can be performed only once.

As the referee pointed out, the algorithm can be interpreted in a most classical way as the Newton-like method where the Inverse-Jacobian approximation is $A^{-1} + \alpha_k \left(E - A^{-1}F'\left(x^k\right)\right)$. In fact, the choice of the line determined by $E - A^{-1}F'\left(x^k\right)$ in the space of matrices allows us to keep the matrix in (2.1) constant. Of course, there are other possibilities for correction of the Inverse-Jacobian, such as correction along $A - F'\left(x^k\right)$ but in this case the unknown matrix A^{-1} appears in the right-hand-side vector. Since the main idea of this paper was to reduce the linear algebra work treating the whole iterative sequence as the sequence of the solutions of the linear systems with the same coefficient matrix and getting the improvement by a perturbation of the right-hand-side vector, the choice of $E - A^{-1}F'\left(x^k\right)$ seems optimal.

Obviously, the new method for $A = F'(x^0)$ and $\alpha_k \equiv 0$ reduces to the fixed Newton method. Although the Newton method is not a special case of MRV, with a good choice of A and α_k we can achieve quite good behavior of the new method.

Let us now state the basic assumptions that will be used for proving local convergence of the method defined by the MRV Algorithm. As usual with the local convergence theories, we assume that there exists $x^* \in R^n$ that solves the problem, that is,

$$F\left(x^*\right) = 0$$

and F' is continuous and nonsingular at x^*.

The following lemmas are used for proving the convergence result.

Lemma 2.1 ([19]). *Let* $F : R^n \to R^n$ *be continuously differentiable at* x^*. *Then for any* $\gamma > 0$ *there exists* $\varepsilon > 0$ *such that*

$$\|F\left(x\right) - F\left(x^*\right) - F'\left(x^*\right)\left(x - x^*\right)\| \le \gamma \|x - x^*\|$$

for $\|x - x^*\| < \varepsilon$.

Lemma 2.2 ([6]). *Let $M \in R^{n \times n}$. If $\|M\| < 1$, then $(E - M)^{-1}$ exists and*

$$(E - M)^{-1} = \sum_{j=0}^{\infty} M^j.$$

Lemma 2.3 ([6]). *Let $M, N \in R^{n \times n}$. If M is nonsingular and $\left\|M^{-1}(M - N)\right\| < 1$, then N is nonsingular and*

$$\|N^{-1}\| \leq \frac{\left\|M^{-1}\right\|}{1 - \left\|M^{-1}(N - M)\right\|}.$$

We are now ready to prove the local convergence result which is the theorem of two neighborhoods.

Theorem 2.4. *Let F satisfy the basic assumptions and $r \in (0,1)$. There exist $\varepsilon = \varepsilon(r)$, $\delta = \delta(r)$ and $\overline{\alpha} > 0$ such that if $\left\|x^0 - x^*\right\| \leq \varepsilon$, $\|A - F'(x^*)\| \leq \delta$ and $|\alpha_k| < \overline{\alpha}$, then the sequence $\left\{x^k\right\}$ generated by Algorithm MRV converges q-linearly to x^* and satisfies*

$$\|x^{k+1} - x^*\| \leq r\|x^k - x^*\|$$

for all $k = 0, 1, \ldots$.

Proof. There exists $\alpha' > 0$ such that for $|\alpha_*| \leq \alpha'$ we have

(2.2) $\|\alpha_* H(x^*)\| < 1.$

The continuity of H allows us to claim

(2.3) $\|\alpha H(x)\| < 1$

for $|\alpha| \leq \alpha'' \leq \alpha'$ and $\|x - x^*\| \leq \varepsilon_1$. Let us denote

$$\Gamma(x, \alpha) = (E - \alpha H(x))^{-1} A.$$

From (2.2), Lemma 2.2 and nonsingularity of A, $\Gamma(x^*, \alpha_*)$ is nonsingular and

$$\left\|\Gamma^{-1}(x^*, \alpha_*)\right\| \leq \beta_1$$

for some $\beta_1 > 0$. Let $\beta = 2\beta_1$. The mapping $\Gamma : R^n \times R \to R^n$ is continuous, so for any $\delta_1 > 0$ we have

(2.4) $\|\Gamma(x, \alpha) - \Gamma(x^*, \alpha_*)\| \leq \delta_1$

for α'', ε_1 small enough. Now Lemma 2.3 implies that $\Gamma(x, \alpha)$ is nonsingular for $\delta_1 < \beta^{-1}$ and

(2.5) $\|\Gamma^{-1}(x, \alpha)\| \leq \beta.$

Denote

(2.6) $\|\alpha_* \Gamma(x^*, \alpha_*) - E\| = C$

and choose $\varepsilon \leq \varepsilon_1$, $\overline{\alpha} \leq \alpha''$, δ such that

(2.7) $\|F(x) - F(x^*) - F'(x^*)(x - x^*)\| \leq \gamma \|x - x^*\|,$

$$\|A - F'(x^*)\| \leq \delta$$

and

$$\beta(C\delta + \delta_1 + \gamma) \leq r < 1$$

for $\|x - x^*\| \leq \varepsilon$ and $|\alpha| \leq \overline{\alpha}$.

Now we prove the statement by induction on k. Let $\|x^k - x^*\| \leq \varepsilon$, $|\alpha_k| \leq \overline{\alpha}$ and $\Gamma_k = \Gamma(x^k, \alpha_k)$, $\Gamma_* = \Gamma(x^*, \alpha_*)$. Then, by (2.4)

$$\|\Gamma_k - \Gamma_*\| \leq \delta_1$$

and

$$
\begin{aligned}
\|x^{k+1} - x^*\| &= \left\| -\Gamma_k^{-1} F\left(x^k\right) + x^k - x^* \right\| \\
&= \left\| \Gamma_k^{-1} \left(\begin{array}{c} \left[\Gamma_k - F'\left(x^*\right)\right]\left(x^k - x^*\right) \\ -\left[F\left(x^k\right) - F\left(x^*\right) - F'\left(x^*\right)\left(x^k - x^*\right)\right] \end{array} \right) \right\| \\
&\leq \|\Gamma_k^{-1}\| \left(\begin{array}{c} \left\|\Gamma_k - F'\left(x^*\right)\right\| \left\|x^k - x^*\right\| \\ +\left\|F\left(x^k\right) - F\left(x^*\right) - F'\left(x^*\right)\left(x^k - x^*\right)\right\| \end{array} \right).
\end{aligned}
$$

Also

$$\|\Gamma_k - F'\left(x^*\right)\| \leq \|\Gamma_k - \Gamma_*\| + \|\Gamma_* - F'\left(x^*\right)\|$$

and

$$
\begin{aligned}
\|\Gamma_* - F'\left(x^*\right)\| &= \left\| \left(E - \alpha_* H\left(x^*\right)\right)^{-1} A - \left(A + H\left(x^*\right)\right) \right\| \\
&= \left\| \left(\sum_{j=0}^{\infty} \left(\alpha_* H\left(x^*\right)\right)^j\right) A - A - H\left(x^*\right) \right\| \\
&= \left\| \alpha_* H\left(x^*\right) \left(\sum_{j=0}^{\infty} \left(\alpha_* H\left(x^*\right)\right)^j\right) A - H\left(x^*\right) \right\| \\
&\leq \|H\left(x^*\right)\| \left\| \alpha_* \left(E - \alpha_* H\left(x^*\right)\right)^{-1} A - E \right\| \\
&\leq \delta C,
\end{aligned}
$$

which follows from (2.2), Lemma 2.2 and (2.6). From these inequalities we obtain

$$
\begin{aligned}
\|x^{k+1} - x^*\| &\leq \beta\left(\delta C + \delta_1 + \gamma\right) \|x^k - x^*\| \\
&\leq r \|x^k - x^*\|
\end{aligned}
$$

and $\|x^{k+1} - x^*\| \leq \varepsilon$, so the inductive step is finished and the statement is proved. $\qquad\square$

3. THE CHOICE OF α_k

In this section we will analyze some strategies for choosing the parameter α. Obviously, the easiest way is to assume

$$(3.1) \qquad \alpha_k \equiv \alpha = const$$

in all iterations.

Let us consider the linear model $M(x)$ of $F(x)$ around the current iterate $x^k \in R^n$

$$(3.2) \qquad M(x^k + d) = F(x^k) + F'(x^k)d.$$

The Newton correction s_N^k is the solution of $M\left(x^k + d\right) = 0$. In the MRV method the correction s^k is

$$s^k = -A^{-1}(E - \alpha_k H(x^k))F(x^k),$$

so the reasonable way to choose the parameter α_k is to solve the following minimization problem:

(3.3) $\alpha_k^{op} = \arg\min_{\alpha \in R} ||M(x^k + s^k)||_2^2,$

i.e.,

(3.4) $\alpha_k^{op} = \arg\min_{\alpha \in R} ||F(x^k) + F'(x^k)(-A^{-1}(E - \alpha H(x^k))F(x^k)||_2^2.$

Using $F'(x^k) = A + H(x^k)$ and with the notation

$$M = H(x^k)A^{-1}, \quad v = MF(x^k), \quad w = H(x^k)F(x^k), \quad t = Mw,$$

(3.4) becomes

$$\alpha_k^{op} = \arg\min_{\alpha \in R} ||v - \alpha w - \alpha t||_2^2.$$

The exact solution of the previous problem is

(3.5) $\alpha_k^{op} = \dfrac{\langle v, w + t \rangle}{\langle w + t, w + t \rangle}.$

This choice is referred to as the optimal parameter. In this case we need to solve two systems of linear equations with the matrix A,

$$Av_1 = F(x^k)$$

and

$$At_1 = w.$$

After that

$$\begin{aligned} v &= H(x^k)v_1, \quad t = H(x^k)t_1, \\ w &= H(x_k)F(x^k), \end{aligned}$$

α_k is given by (3.5) and

$$s^k = -(v_1 - \alpha_k t_1).$$

One iteration requires $\mathcal{O}(2n^2)$ flops for v_1 and t_1 and three matrix-vector products. Thus, one iteration requires the double work in comparison with (3.1) but numerical examples in Section 4 show that the effort is justified.

We are now going to prove the convergence theorem for the MRV method with the optimal parameter assuming that the basic assumptions from Section 2 are satisfied. The proof is based on the following theorem from [4].

Theorem 3.1. *Assume that* $\eta_k \leq \eta < 1$. *There exists* $\varepsilon > 0$ *such that, if* $||x - x^*|| \leq \varepsilon$, *then the sequence of inexact Newton iterates* $\{x^k\}$ *converges to* x^*. *Moreover, the convergence is linear in the sense that*

$$||x^{k+1} - x^*||_* \leq r||x^k - x^*||_*,$$

where $||y||_* = ||F'(x^*)y||.$

In order to prove the convergence of MRV method we are going to consider it as an Inexact Newton method.

Theorem 3.2. *There exist* $\delta, \varepsilon > 0$ *such that, if* $||x - x^*|| \leq \varepsilon$ *and* $||A - F'(x^*)|| \leq \delta$, *then the sequence* $\{x^k\}$ *generated by Algorithm MRV with the parameters* $\{\alpha_k\}$ *given in* (3.5) *converges to* x^*. *The convergence is linear in the sense that*

$$||x^{k+1} - x^*||_* \leq r||x^k - x^*||_*.$$

Proof. Choose $\varepsilon, \delta > 0$ such that $||A - F'(x)|| \leq 2\delta$, for $||x - x^*|| \leq \varepsilon$ and $\delta < \eta/(2M)$, for $\eta < r < 1$ and $||A^{-1}|| \leq M$. Let us define functions

$$s(x, \alpha) = -A^{-1}(E - \alpha H(x)) F(x)$$

and

$$\eta(x, \alpha) = ||F(x) + F'(x) s(x, \alpha)||_2.$$

Since

$$\eta(x, \alpha^{op}) \leq \eta(x, 0)$$

and

$$
\begin{aligned}
\eta(x, 0) &= ||F(x) - F'(x) A^{-1} F(x)|| \\
&\leq ||E - F'(x) A^{-1}|| \, ||F(x)|| \\
&\leq ||A - F'(x)|| \, ||A^{-1}|| \, ||F(x)|| \\
&\leq 2\delta M ||F(x)|| \\
&\leq \eta ||F(x)||,
\end{aligned}
$$

considering the MRV method as an Inexact Newton method with

$$||r^k|| = \eta\left(x^k, \alpha_k^{op}\right),$$

we have

$$\eta\left(x^k, \alpha_k^{op}\right) \leq \eta\left(x^k, 0\right) \leq \eta ||F(x)||$$

and the statement follows from Theorem 3.1. $\square$

Let us mention that for some special systems the correction s^k for α_k^{op} becomes the solution of $M(x^k + d) = 0$, in other words $s^k = s_N^k$. If for every k we can determine $\alpha_k = \alpha_k^{op}$ such that $M(x^k + s^k) = 0$, the MRV iterative sequence will be the Newton sequence. This is true for the systems of the form $F(x) = \left[\begin{smallmatrix} F_1(x) \\ Cx \end{smallmatrix}\right]$, where $C \in R^{(n-1) \times n}$ and $F_1 : R^n \to R$ is a nonlinear mapping.

4. NUMERICAL RESULTS

In this section we present the results of numerical experiments. Let us describe first the algorithms we used.

1. MRV denotes the method proposed in this paper with the optimal parameter α_k, (3.5) and $A = F'(x^0)$.
2. MRVF stands for the same method with a constant value of the parameter, $\alpha_k = \alpha$, $A = F'(x^0)$.
3. The Newton method (N).
4. The Fixed Newton method (FN), as described in Algorithm FN.
5. The Bogle-Perkins method (BP).
6. The Schubert method (S).
7. The row scaling method (RS), with restart after every 5 iterations.

The Bogle-Perkins method was introduced in [1], the Shubert method was introduced in [20] and the row-scaling method was introduced in [11].

In the examples which follow we say that the method converges if the conditions

$$||x^k - x^{k+1}|| \leq 10^{-4} ||x^k|| + 10^{-4} \quad \text{and} \quad ||F(x^k)|| \leq 10^{-4}$$

are satisfied. If $\left\| F\left(x^k\right)\right\| > M_{\max}$ holds, we consider the method divergent and if after $I_{\max}$ iterates neither convergence nor divergence conditions are satisfied, we consider the method slowly convergent. Both cases are marked with $*$. We used $M_{\max} = 10^{10}$ and $I_{\max} = 100$.

Let us now describe the nonlinear systems used in our numerical study.

Problem 4.1. See [11].

$$F_1(x) = -2x_1^2 + 3x_1 - 2x_2 + 0.5x_{\alpha_1} + 1,$$
$$F_i(x) = -2x_i^2 + 3x_i - x_{i-1} - 2x_{i+1} + 0.5x_{\alpha_i} + 1, \quad i = 2, \dots, n-1,$$
$$F_n(x) = -2x_n^2 + 3x_n - x_{n-1} + 0.5x_{\alpha_n} + 1$$

for α_i, $i = 1, \dots, n$, randomly chosen in the interval

$$\alpha_i \in \{\max\{1, i-p\}, \min\{n, i+p\}\}$$

and parameter p which defines the bandwidth. We used two start approximations: $x^0 = (-1, \dots, -1)^T$ and $x^0 = (-3, \dots, -3)^T$.

The Jacobian is treated as a band matrix with upper and lower bandwidth p, so the cost of LU factorization is $\mathcal{O}(np^2)$ and the additional $\mathcal{O}(2n(p+1))$ flops for the solution of linear system (see Golub, van Loan [10]). The matrix $H(x)$ is a diagonal matrix in this case.

Problem 4.2. Structured Jacobian problem (see [14]).

$$
\begin{aligned}
F_1(x) &= -2x_1^2 + 3x_1 - 2x_2 + 3x_{n-4} - x_{n-3} - x_{n-2} + 0.5x_{n-1} - x_n + 1, \\
F_i(x) &= -2x_i^2 + 3x_i - x_{i-1} - 2x_{i+1} + 3x_{n-4} - x_{n-3} - x_{n-2} \\
&\quad + 0.5x_{n-1} - x_n + 1, \qquad\qquad\qquad\qquad i = 2, \dots n-1, \\
F_n(x) &= -2x_n^2 + 3x_n - x_{n-1} + 3x_{n-4} - x_{n-3} - x_{n-2} + 0.5x_{n-1} - x_n + 1,
\end{aligned}
$$

with $x_i^0 = -1$, $i = 1, \dots, n$.

Problem 4.3. Generalization of Brown (see [3]).

$$F_1(x) = -1 + \prod_{i=1}^{n} x_i,$$
$$F_i(x) = -(n+1) + x_i^p + \sum_{k=1}^{n} x_k, \quad i = 2, \dots, n,$$

with the initial approximation $x^0 = (0.9, \dots, 0.9)^T$.

For $p = 1$ we obtain the function from [2]. In this case nonlinearity appears only in the first equation. For $p > 1$ the system becomes more nonlinear. This system is ill-conditioned. The condition number of the Jacobian at the initial point increases with increasing order due to the values of the Jacobian elements in the first row.

Problem 4.4. Band Broyden, [11].

$$F_i(x) = (3 + 5x_i^2)x_i + 1 - \sum_{j \in I_i}(x_j + x_j^2), \quad i = 1, \dots, n,$$

where $I_i = \{\max\{1, i-p\}, \dots, \min\{n, i+p\}\} - \{i\}$. The initial approximation was $(-2, \dots, -2)^T$. For this function the Jacobian is a band matrix with lower and upper bandwidth p, so the cost is the same as for Problem 1.

Problem 4.5. Chandrasekhar H-equation, [17].

The integral equation

$$F(H)(x) = H(x) - \left(1 - \frac{C}{2} \int_0^1 \frac{xH(y)dy}{x+y}\right)^{-1} = 0$$

is discretized using the midpoint quadrature formula

$$\int_0^1 f(t)dt = \frac{1}{n} \sum_{j=1}^n f(t_j),$$

for $t_j = (j - \frac{1}{2})h$, $h = \frac{1}{n}$, $1 \le j \le n$, and the following system is obtained:

$$F_i(x) = x_i - \left(1 - \frac{c}{2n} \sum_{j=1}^n \frac{t_i x_j}{t_i + t_j}\right)^{-1}, \quad i = 1, \dots, n.$$

The integral equation and the discrete analog have two solutions for $c \in (0,1)$, but only one of them has physical meaning. The problem becomes more difficult as c approaches 1. We report results for $c = 0.9$, 0.99, 0.9999. For $c = 0.9999$ the Jacobian at the solution is nearly singular. In all the cases the Jacobian is dense, so the solution of the Newton system with LU factorization requires $\mathcal{O}(n^3)$ flops. All tests were performed using $(1, \dots, 1)^T$ as the initial approximation.

Problem 4.6. Singular Broyden, [11].

$$F_1(x) = ((3 - hx_1)x_1 - 2x_2 + 1)^2,$$
$$F_i(x) = ((3 - hx_i)x_i - x_{i-1} - 2x_{i+1} + 1)^2,$$
$$F_n(x) = ((3 - hx_n)x_n - x_{n-1} + 1)^2.$$

We used $x^0 = (-1, \dots, -1)^T$ as a starting approximation and $h = 2$.

Problem 4.7. Trigonometric System, [14].

For $k = \mathrm{div}\,(i - 1, 5)$,

$$F_i(x) = 5 - (k+1)(1 - \cos x_i) - \sin x_i - \sum_{j=5k+1}^{5k+5} \cos x_j, \quad i = 1, \dots, n,$$

with $x_i^0 = 1/n$.

Problem 4.8. Trigonometric-Exponential System, [14].

$$
\begin{aligned}
F_1(x) &= 3x_1^2 + 2x_2 - 5 + \sin(x_1 - x_2)\sin(x_1 + x_2), \\
F_i(x) &= 3x_i^2 + 2x_{i+1} - 5 + \sin(x_i - x_{i+1})\sin(x_i + x_{i+1}) + 4x_i \\
&\quad -x_{i-1}\exp(x_{i-1} - x_i) - 3, \qquad\qquad i = 2, \dots, n-1, \\
F_n(x) &= 4x_n - x_{n-1}\exp(x_{n-1} - x_n) - 3,
\end{aligned}
$$

with $x_i^0 = 0$.

Problem 4.9. Generalized function of Rosenbrock, [14].

$$
\begin{aligned}
F_1(x) &= -4c\left(x_2 - x_1^2\right)x_1 - 2(1 - x_1), \\
F_i(x) &= 2c\left(x_i - x_{i-1}^2\right) - 4c\left(x_{i+1} - x_i^2\right)x_i - 2(1 - x_i), \quad i = 2, \dots, n-1, \\
F_n(x) &= 2c\left(x_n - x_{n-1}^2\right),
\end{aligned}
$$

with $c = 2$ and $x_i^0 = 1.2$.

TABLE 1.

Pr.	x^0	Paramet.	N	FN	MSV	MSVF	S	BP	RS
1	$x_i^0 = -1$	$p = 5$	4	12	6	6 ($\alpha = -0.25$)	6	5	6
1	$x_i^0 = -3$	$p = 5$	6	47	10	12 ($\alpha = -0.2$)	12	8	7
2	$x_i^0 = -1$		5	16	14	8 ($\alpha = -0.3$)	8	7	7
4	$x_i^0 = -2$	$p = 12$	6	$*$	14	14 ($\alpha = -0.05$)	25	$*$	7
4	$x_i^0 = -2$	$p = 30$	6	$*$	18	38 ($\alpha = -0.01$)	21	8	7
5	$x_i^0 = 1$	$c = 0.9$	4	7	4	4 ($\alpha = -1.8$)	4	6	6
5	$x_i^0 = 1$	$c = 0.99$	5	21	5	4 ($\alpha = -4.5$)	6	27	8
5	$x_i^0 = 1$	$c = 0.9999$	8	$*$	8	30 ($\alpha = -5$)	10	$*$	17
6	$x_i^0 = -1$		13	$*$	18	$*$ ($\alpha = 0.05$)	$*$	$*$	17
9	$x_i^0 = 1.2$		4	18	9	9 ($\alpha = -0.08$)	7	6	$*$

TABLE 2.

	N	FN	MRV	MRVF	S	BP	RS
R	1	0.6	1	0.9	0.9	0.7	0.9
E	0.98	0.2929	0.6552	0.565	0.595	0.6475	0.706
$E \times R$	0.98	0.17574	0.6552	0.5085	0.5355	0.45325	0.6354

The numbers of iterations for the following ten problems are presented in Table 1. The third column contains the parameters of problems. The parameter α used in the MRVF method is presented in column MSVF. In these tests we used $n = 100$.

Three indices are used to collect the data for comparison (see [1]). If r_{ij} is the number of iterations required to solve the problem i by the method j, $r_{ib} = \min_j r_{ij}$, i.e., the best result for problem i by any of tested methods, t_j the number of successes by method j and n_j the number of problems attempted by method j, then the robustness index is

$$R_j = \frac{t_j}{n_j},$$

the efficiency index is

$$E_j = \sum_{i=1, r_{ij} \neq 0}^{m} \left(\frac{r_{ib}}{r_{ij}} \right) / t_j,$$

and the combined robustness and efficiency index:

$$E_j \times R_j = \sum_{i=1, r_{ij} \neq 0}^{m} \left(\frac{r_{ib}}{r_{ij}} \right) / n_j.$$

R is the percentage of cases in which each method found a solution. For E and $E \times R$ the best possible result is 1, and larger values of indices indicate a better result. The results from Table 1 are summarized in Table 2. The new methods appear to be fairly competitive with the considered Newton-like methods.

We have also done some tests on systems with larger dimension, $n = 1000$. The results are reported in Table 3.

The problem 3 has a nearly singular Jacobian. The results are reported in Table 4.

TABLE 3.

Pr.	x^0	Param.	N	FN	MRV	MRVF	S	BP	RS
1	$x_i^0 = -1$	$p = 5$	4	13	6	$6\,(\alpha = -0.25)$	7	5	6
1	$x_i^0 = -3$	$p = 5$	6	*	10	$11\,(\alpha = -0.1)$	13	8	7
2	$x_i^0 = -1$		5	17	*	$8\,(\alpha = -0.3)$	8	7	7
7	$x_i^0 = \frac{1}{n}$		5	171	8	$15\,(\alpha = 1)$	*	*	*
8	$x_i^0 = 0$		5	*	12	$15\,(\alpha = -0.08)$	11	10	*

TABLE 4. Problem 3, $x_i^0 = 0.9$

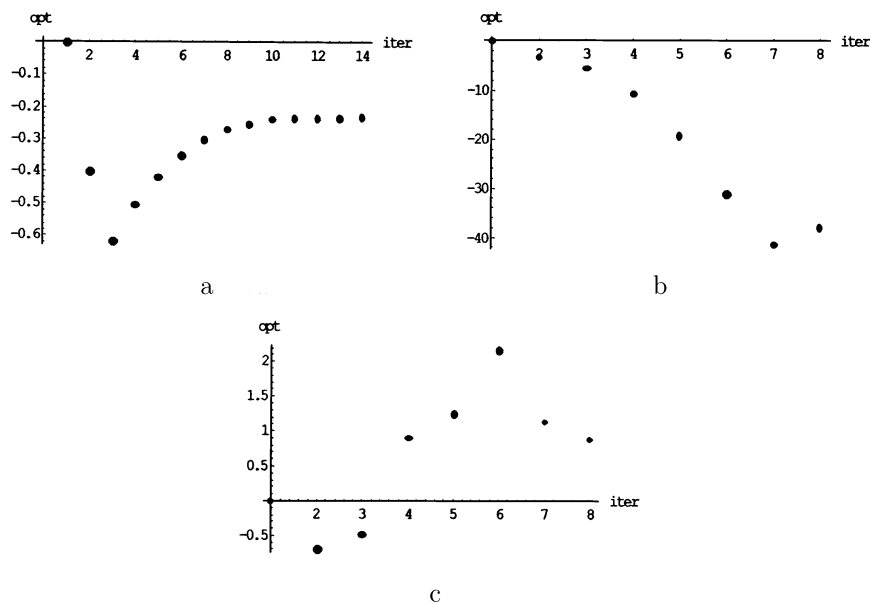| n	Parameters	N	FN	MRV	MRVF
5	$p = 1$	13	*	13	$11\ (\alpha = -0.5)$
7	$p = 1$	14	*	14	$*\ (\alpha = -0.5)$
9	$p = 1$	14	*	14	$*\ (\alpha = -0.4)$
5	$p = 2$	13	14	13	$9\ (\alpha = -0.3)$
7	$p = 2$	13	*	14	$14\ (\alpha = -0.4)$
9	$p = 2$	14	*	14	$*\ (\alpha = -0.4)$
5	$p = 3$	13	11	13	$8\ (\alpha = -0.2)$
7	$p = 3$	13	*	14	$17\ (\alpha = -0.4)$
9	$p = 3$	14	*	15	$*\ (\alpha = -0.4)$

FIGURE 1.

The optimal choice for the parameter α_k is defined in Section 3. Obviously, the parameter has a strong influence on the behavior of the method. Figures 1a–1c show the values of the optimal parameter for Problem 3 ($n = 7$, $p = 2$), Problem 5 ($n = 100$, $c = 0.999$) and Problem 8 ($n = 100$, $x_i^0 = 1/i$), respectively.
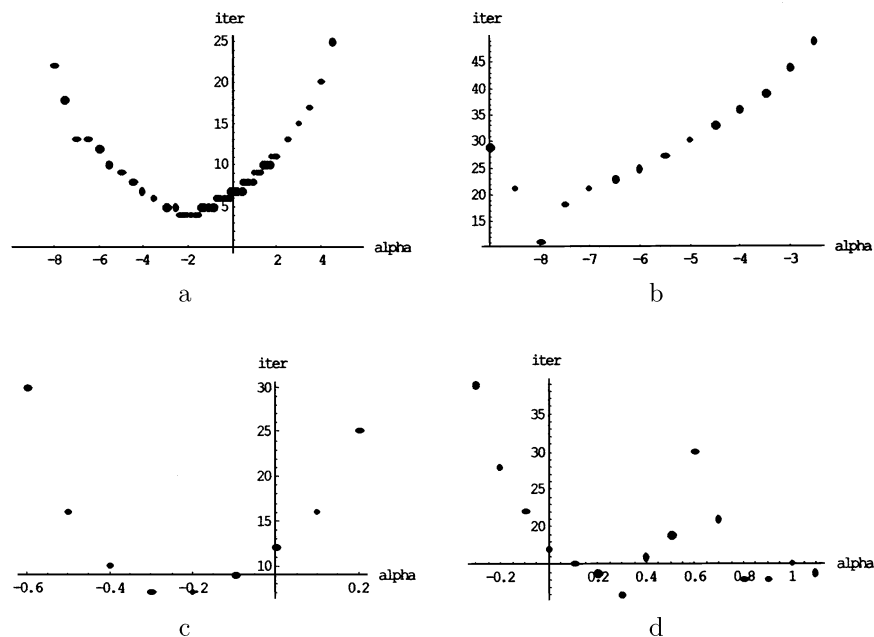
FIGURE 2.

As we mentioned before, the most simple choice is $\alpha_k = const$. Figures 2a–2d show the number of iterations for different values of $\alpha_k = const$, for Problem 5 ($n = 100$, $c = 0.9$), Problem 5 ($n = 100$, $c = 0.999$), Problem 1 ($n = 100$, $p = 5$) and Problem 7 ($n = 100$) respectively. We can see that the FN method ($\alpha = 0$) is not the best one in any of the problems considered. Obviously, the best choice for a constant parameter depends on the problem we are solving. One of the possibilities for the choice of a constant parameter is to calculate the optimal parameter only in the first iteration and keep that value to the end of the process,

$$\alpha = \alpha_1^{opt}.$$

This choice happens to be satisfactory with the problems considered here, although this conclusion is influenced by the collection of test problems.

5. CONCLUSIONS

We presented a new method for solving nonsingular systems of equations. The method belongs to the class of Newton-like methods.

The central idea was to keep the initial approximation of the Jacobian fixed (and its LU factorization) and to modify the right-hand-side vector in each iteration in order to make it computationally cheaper than the Newton method and faster than the fixed Newton method. The modification is obtained using a relaxation parameter and assuming that the Jacobian is available in each iteration. Some suitable choices for the relaxation parameter are proposed. Numerical experiments indicate that the parameter α plays an important role in the proposed algorithm and that the question needs further investigation, both from the theoretical and computational points of view.

As we have already mentioned, the relaxation parameter has a crucial influence on the behavior of the MRV method. One possible generalization would be to consider relaxation with a diagonal matrix $(diag\,(\alpha_1,\dots,\alpha_n))$ instead of only one parameter. Of course, in this way we can obtain the Newton method as a special case, using optimal values of α_i, but calculation of n parameters is expensive in the general case. The idea seems promising with $\alpha_i = 1$ except in a few rows for solving systems that are ill-conditioned, mainly because of instability in a few components of the function. This is the case with the systems which arise from the discretization of singularly perturbed boundary value problems. Since nonequidistant discretization grids are a necessity in these problems, the change of the step size usually implies more problems at some specific components of the function (see [13], [9]). Thus, the idea would be to use suitable relaxation parameters and modify only those components of the right-hand-side vector.

Acknowledgment

References

[1] I. D. L. Bogle, J. D. Perkins, *A new sparsity preserving quasi-Newton update for solving nonlinear equations,* SIAM J. Sci. Statist. Comp., **11** (1990), pp. 621-630. MR **91f:**65099

[2] K. M. Brown, *A quadratically convergent Newton-like method based upon Gaussian elimination,* SIAM J. Numer. Anal., **6** (1969), pp.560-569. MR **41:**7834

[3] J. C. P. Bus, *Numerical solution of systems of nonlinear equations,* Tract 122, Mathematisch Centrum, Amsterdam, 1980. MR **81m:**65072

[4] R. S. Dembo, S. C. Eisenstat and T. Steihaug, *Inexact Newton methods,* SIAM J. Numer. Anal., **19** (1982), pp.400-408. MR **83b:**65056

[5] J. E. Dennis and E. S. Marwil, *Direct secant updates of matrix factorizations,* Math. Comp., **38** (1982), pp.459-474. MR **83d:**65159

[6] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations,* Prentice Hall, Englewood Cliffs, NJ, 1983. MR **85j:**65001

[7] L. C. W. Dixon, *On the impact of automatic differentiation on the relative performance of parallel truncated Newton and variable metric algorithms,* SIAM J. Optim. **1** (1991), 475-486. MR **93c:**90079

[8] S. C. Eisenstat and H. F. Walker, *Choosing the forcing terms in an inexact Newton method,* SIAM J. Sci. Comput., **17** (1996), pp.16-32. MR **96k:**65037

[9] P.A. Farrell, J.J.H. Miller, E. O'Riordan, G. I. Shishkin, *A uniformly convergent finite difference scheme for a singularly perturbed semilinear equation,* SIAM J. Numer. Anal. **33**(3), (1996), 1135-1149. MR **97b:**65086

[10] G. H. Golub, C. F. van Loan, *Matrix Computations,* The Johns Hopkins University Press, Baltimore, MD, 1983. MR **85h:**65063

[11] M. A. Gomes-Ruggiero, J. M. Martínez and A. C. Moretti, *Comparing algorithms for solving sparse nonlinear system of equations,* SIAM J. Sci. Comput., **13** (1992), pp.459-483. MR **92j:**65077

[12] D. Herceg, N. Krejić, Z. Lužanin, *Quasi-Newton's method with correction,* Novi Sad J. Math., **26**(1), 1996, pp. 115-127. MR **98h:**65021

[13] D. Herceg, N. Krejić, *On a numerical method for discrete analogues of boundary value problems,* Nonlinear Analysis, Theory, Methods & Applications, **30**,1 (1997), 9-15. CMP 98:06

[14] L. Lukšan, *Inexact trust region method for large sparse systems of nonlinear equations,* JOTA, **81**(3), 1994, pp. 569-590. MR **95b:**65072

[15] L. Lukšan, J. Vlček, *Computational experience with globally convergent descent methods for large sparse systems of nonlinear equations,* Optimization Methods and Software **8** (1998), 201-224. MR **99a:**65071

[16] Z. Lužanin, *Global and local convergence of modifications of Newton method,* Ph.D. Thesis, Institute of Mathematics, Faculty of Science, University of Novi Sad, 1997.

[17] C. T. Kelley, *Iterative methods for linear and nonlinear equations,* SIAM, Philadelphia, 1995. MR **96d:**65002

[18] 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), pp.1034-1049. MR **91h:**65084

[19] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables,* Academic Press, New York, 1970. MR **42:**8686

[20] I. K. Schubert, *Modification of a quasi-Newton method for nonlinear equation with a sparse Jacobian,* Math. Comp. **24** (1970), pp.27-30. MR **41:**2923

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

   E-mail address: natasa@unsim.im.ns.ac.yu

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

   E-mail address: luzanin@uns.ns.ac.yu