Using a new line search method with gradient direction to solve nonlinear systems of equations

Karrar Habeeb Hashim¹, Mushtak A.K. Shiker²

Department of Mathematics, College of Science, Mustansiriyah University

teacher89karrar@gmail.com
mmttmhh@yahoo.com

Abstract. The line search techniques together with the Newton method are the best methods to solve nonlinear systems of equations. These methods use the gradient directions because they required low storage. In this paper, we suggest a new line search algorithm with gradient direction to solve the nonlinear systems of equations. The purpose of this algorithm is to reduce the number of iterations and the function evaluations, and it can increase the effectiveness of the approach. The algorithms global convergence is proved. The numerical results indicates the efficiency of the new algorithm and it is promised for solving nonlinear systems of equations.

Keywords. Line search technique, Nonlinear systems, Gradient direction, Global convergence.

1. Introduction

Some problems arise in the various fields of computational science and geometry, among the problems that arise in these areas, especially in the interpretation of nonlinear partial differential equations and the problem of specific value, etc, are nonlinear systems.

To solve these nonlinear systems, many researchers have proposed some common algorithms that provide appropriate ways and means to solve these problems.

Among those problems that will be addressed are optimization problems and since the convergence and the work of these problems are slow so we will use the monotone strategy to solve it.

Consider the nonlinear system of equations

\[ F(x) = 0 \]  

Where \( F: R^n \rightarrow R^n \) is continuous and monotone, i.e.

\[ \langle F(x) - F(y), x - y \rangle \geq 0, \forall x, y \in R^n. \]

In different applications, such as sub-problems in general algorithms, this type of nonlinear monotone equations is created [1], also through a fixed point map or a natural map, many monotonous variational inequality can be transformed into monotonous nonlinear equations [2].

Newton's method, the methods of quasi-Newton and their different forms are important to solve smooth systems of equations because of the rates of rapid convergence of local [3,4], also the line search technique used for some merit functions to ensure global convergence [5, 6].

In (1986), Griewank calculated the global convergence results of nonlinear equations for Broyden’s rank one form [7]. In (1998), and to solve the systems of monotonous equations, Svaiter and Solodov introduced a Newton-type algorithm, and by using the hybrid projection method they showed that this method converges globally [8].
In (1999), and to solve symmetric nonlinear equations Li and Fukushima presented a Gauss–Newton-based BFGS method and prove their global convergence [9]. Because of the need for these methods to calculate and store the matrix then it will be inappropriate in large-scale nonlinear equations [10]. In order to get rid of this defect, Nocedal proposed a BFGS method of limited memory (L-BFGS) for the problems of unconstrained minimization [11], because of the low storage of this method, numerical results showed that this approach is very competitive [12].

Many methods and algorithms have also been proposed to solve non smooth systems of equations, such as semi smooth Newton method [13] and trust region methods [14, 15], and since these methods converge easily from any initial guess that is strong enough, then they are more attractive than others. Despite its attractiveness, it contains some negatives, which are important in large values of \( n \) and they need to solve a linear system of equations in each iteration by using or approximating the Jacobian matrix.

In (1988), both Braz and Bruan introduced the spectral gradient method, which is considered a low-cost non-monotonous schemes used to find local minimizers [16], then in (2003) Cruz and Raydan expanded this method to solve nonlinear equations, also by using some merit functions, Cruz and Raydan have made some attempts to convert nonlinear equations into a unconstrained minimization problem [17].

In (2006), in order to address unconstrained nonlinear monotone equations, Zhang and Zhou introduced some important and effective adjustments to the ways listed in the [18] by adopting a new line search strategy with a new step length with taking into account the monotony of \( F \) [19].

In this paper, and without using any descent method and any merit function, we suggest a new algorithm to solve (1), and by using some assumptions we will prove that it is a well-define and also globally converges to solutions of (1), then we will see how efficient the proposed new method is compared to method in [19] and SCGP method in [20] by using numerical results.

Now, we will give our algorithm.

2. Algorithm (KA):
Step 0. Given an initial point \( x_0 \in \mathbb{R}^n \), positive constant \( m > 0 \) and constants \( \beta \in (0,1) \), \( \rho, \delta \in [0.5,1) \).
Set \( k := 0 \).
Step 1. Determine the search direction \( d_k \) by
\[
d_k = \begin{cases} 
-F(x_k) & \text{if } k = 0, \\
-\lambda_k F(x_k) & \text{otherwise},
\end{cases}
\]
where \( \lambda_k = \frac{s_k}{y_k s_k} \), \( s_k = x_{k+1} - x_k \), with \( y_k = F(x_{k+1}) - F(x_k) + m_k s_k \).
Where \( m_k = \frac{1}{m \| d_k \|^2} \).
Stop if \( d_k = 0 \).
Step 2. If \( \| F(x_k) \| = 0 \), then stop.
Step 3. Determine step length \( \alpha_k = \beta^w k \) such that \( w_k \) is the smallest non-negative integer \( w \) satisfying:
\[
-\langle F(x_k + \beta^w d_k), d_k \rangle \geq \rho(1 - \delta_k)\beta^w \| d_k \|^2
\]
where \( \delta_k = \frac{\delta}{1 + \| F(x_k) \|} \), and find the trail point \( z_k = x_k + \alpha_k d_k \).
Step 4. Compute
\[
x_{k+1} = x_k - \frac{(F(x_k), x_k, z_k)}{\| F(x_k) \|^2} F(z_k)
\]
Step 5. Set \( k = k + 1 \) and go to step 1.

Assumptions:
\( (A_1) \) the solution of (1) is non-empty.
\( (A_2) \) the mapping \( F \) is monotone and its Lipschitz continuous (LC), that is to say, there is a positive constant \( L > 0 \) such that
\[ \|F(x) - F(y)\| \geq L \|x - y\|, \forall x, y \in \mathbb{R}^n. \]

Remarks. (i) by the monotonicity of \( F \), then in step (1) we have
\[
\begin{align*}
y_k^T s_k &= (F(x_{k+1}) - F(x_k) + m_k s_k) s_k \\
&= (F(x_{k+1}) - F(x_k)) s_k + m_k s_k^T s_k \\
&= (F(x_{k+1}) - F(x_k), x_{k+1} - x_k) + m_k s_k^T s_k \geq m_k s_k^T s_k > 0
\end{align*}
\]
In addition since \( F \) is LC then we also have
\[
\begin{align*}
y_k^T s_k &= (F(x_{k+1}) - F(x_k) + m_k s_k) s_k \\
&= (F(x_{k+1}) - F(x_k)) s_k + m_k s_k^T s_k \\
&= (F(x_{k+1}) - F(x_k), x_{k+1} - x_k) + m_k s_k^T s_k \leq (L + m_k) s_k^T s_k
\end{align*}
\]
So from (5) and (6) we have
\[
(L + m_k) s_k^T s_k \geq m_k s_k^T s_k \Rightarrow (L + m_k) \geq m_k
\]
\[
\Rightarrow \frac{F(x_k)}{L + m_k} \leq \|d_k\| \leq \frac{F(x_k) - F(x_k)}{m_k}
\]
\[(7)\]
(ii) the algorithm (KA) generates an infinite sequence \( \{x_k\} \) if we assume that \( F(x_k) \neq 0 \) for \( k \geq 0 \).

Now we will introduce two ways to prove that the line search (3) is a well-defined.

First: Suppose that for any nonnegative integer \( w \) and for some iteration index \( k \), the line search (3) is not satisfied i.e.,
\[
-F(x_k + \beta^w d_k), d_k < \rho (1 - \delta_k) \beta^w \| d_k \|^2
\]
Now if we take \( \lim_{w \to \infty} \) for two sides to (8),
\[
\begin{align*}
&-\lim_{w \to \infty} (F(x_k + \beta^w d_k), d_k) < \lim_{w \to \infty} \rho (1 - \delta_k) \beta^w \| d_k \|^2 \\
&\Rightarrow -\lim_{w \to \infty} (F(x_k), d_k) < 0 \\
&\Rightarrow -\lim_{w \to \infty} (F(x_k), d_k) < 0 \\
&\Rightarrow -\lim_{w \to \infty} (F(x_k), d_k) < 0 \\
&\Rightarrow \alpha_k \|F(z_k + d_k, d_k)\| < 0
\end{align*}
\]
Then we have a contradiction, since it is not possible to have each of \( \alpha_k \), \( \|F(z_k)\| \) and \( \|d_k\|^2 \) less than zero, so the line search (3) is well-defined.

Second: By the following lemma

Lemma 2.1. Suppose that \( F \) is LC, and for \( k \geq 0 \), then there is a non-negative number \( w_k \) satisfying
\[
-(F(x_k + \beta^w d_k), d_k) \geq \rho (1 - \delta_k) \beta^w \| d_k \|^2
\]

Proof. assume that there exists \( k_0 > 0 \) such that (2.8) is not satisfied for any nonnegative number \( w \) i.e.,
\[
-(F(x_k + \beta^w d_k), d_k) < \rho (1 - \delta_k) \beta^w \| d_k \|^2
\]
Let \( w \to \infty \) and using the \( F \) yields continuity
\[
-(F(x_k), d_k) \leq 0
\]
From steps (1) and (2) we have \( F(x_k) \neq 0, d_k \neq 0, \forall k \geq 0 \)
Thus from (6) and the value of \( \lambda_k \) we obtain
\[
\lambda_k = \frac{s_k^T s_k}{u_k^T s_k} \geq \frac{s_k^T s_k}{(L + m_k) s_k^T s_k} = \frac{1}{(L + m_k)} > 0 \quad \forall k > 0
\]
This means
\[
-F(x_k), d_k = \lambda_k (F(x_k), F(x_k)) > 0 \\
-F(x_k), d_k = \lambda_k (F(x_k), F(x_k)) = \lambda_k (F(x_k), F(x_k)) > 0 \quad \forall k > 0
\]
Note that the last inequality contradicts with (8), thus the statement is proved.

3. Convergence property
We need the following lemma if we have to obtain global convergence.
**Lemma 3.1** [8] Let $F$ be monotone and $x, y \in \mathbb{R}^n$ satisfy $(F(y), x - y) > 0$. Let

$$x^+ = x - \frac{(F(y), x - y)}{||F(y)||^2}F(y)$$

Then for any $\tilde{x} \in \mathbb{R}^n$ such that $F(\tilde{x}) = 0$, it holds that

$$||x^+ - \tilde{x}||^2 \leq ||x - \tilde{x}||^2 - ||x^+ - x||^2$$

Now we can state our convergence result.

**Theorem 3.2.** Suppose that the algorithm $(KA)$ generated a sequence $\{x_k\}$ and let $F$ is (LC) and monotone such that the solution set of (1) is nonempty. Then for any $\bar{x}$ satisfying $F(\bar{x}) = 0$, we have:

$$||x_{k+1} - \bar{x}||^2 \leq ||x_k - \bar{x}||^2 - ||x_{k+1} - x_k||^2$$

Particularly, the sequence $\{x_k\}$ is bounded. Also, the satisfaction is either that $\{x_k\}$ is finite and the last iterate is a solution, or that the sequence is infinite and $||x_{k+1} - x_k|| = 0$.

**Proof.** First if the algorithm finishes at some iteration $k$, then we have $d_k = 0$ and $F(x_k) = 0$, in this case $x_k$ will be a solution of (1).

Now suppose that $d_k \neq 0$, and $F(x_k) \neq 0$ for all $k$, Then we will get an infinite sequence $\{x_k\}$.

It follows from (3) that

$$\langle F(z_k), x_k - z_k \rangle = \langle F(z_k), x_k - x_k - \alpha_k d_k \rangle$$

(since $z_k = x_k + \alpha_k d_k$)

$$\geq -\alpha_k F(z_k, d_k)$$

$$\geq \rho(1 - \delta_k)\alpha_k^2 ||d_k||^2 = \rho(1 - \delta_k) ||x_k - z_k||^2 > 0$$

Then,

$$\langle F(z_k), x_k - z_k \rangle \geq \rho(1 - \delta_k)\alpha_k^2 ||d_k||^2 = \rho(1 - \delta_k) ||x_k - z_k||^2 > 0 \quad (10)$$

Let $\tilde{x}$ be an any point such that $F(\tilde{x}) = 0$. From lemma 3.1 and (4), (10), we obtain

$$||x_{k+1} - \tilde{x}||^2 \leq ||x_k - \tilde{x}||^2 - ||x_{k+1} - x_k||^2. \quad (11)$$

This mean, the sequence $\{||x_k - \tilde{x}||\}$ is decreasing and hence convergent. Consequently, the sequence $\{x_k\}$ will be bounded, and also we have

$$||x_{k+1} - x_k|| = 0 \quad (12)$$

By (7) It is clear that $\{d_k\}$ holds to be bounded and so is $\{z_k\}$.

Now since $F$ is continuous, then there exist a positive constant $t > 0$ such that

$$||F(z_k)|| \leq t$$

From (4) and (10) we obtain

$$x_{k+1} = x_k - \frac{\langle F(x_k + \beta_k d_k), d_k \rangle}{||F(x_k)||^2}F(z_k)$$

Since

$$\langle F(z_k), x_k - z_k \rangle = -\alpha_k F(z_k, d_k), \text{ then}$$

$$x_{k+1} - x_k = \frac{\alpha_k F(z_k, d_k)}{||F(x_k)||^2}F(z_k) \geq \frac{\rho(1 - \delta_k)\alpha_k^2 ||d_k||^2}{||F(x_k)||^2}F(z_k)$$

So, we have

$$||x_{k+1} - x_k|| \geq \frac{\rho}{t}(1 - \delta_k)\alpha_k^2 ||d_k||^2$$

By (12) and the above inequality we have

$$\lim_{k \to \infty} \alpha_k ||d_k|| = \lim_{k \to \infty} ||x_k - z_k|| = 0 \quad (13)$$

From (7), we get $\liminf_{k \to \infty} ||F(x_k)|| = 0,$ If $\liminf_{k \to \infty} ||d_k|| = 0$.

Since $F$ is continuous and $\{x_k\}$ is bounded, then we have $\{x_k\}$ has some accumulation point $\tilde{x}$ such that $F(\tilde{x}) = 0$. From (11) the converges of sequence $\{||x_k - \tilde{x}||\}$ is holds, also we have $\{x_k\}$ converges to $\tilde{x}$ because $\tilde{x}$ is accumulation point of $\{x_k\}$.

From (7) we get $\liminf_{k \to \infty} ||F(x_k)|| > 0$, If $\liminf_{k \to \infty} ||d_k|| > 0$. Then by (13) we get

$$\lim_{k \to \infty} \alpha_k = 0 \quad (14)$$

(3) gives us

$$-\langle F(x_k + \beta_{w_k} d_k), d_k \rangle < \rho(1 - \delta_k)\beta_{w_k} ||d_k||^2 \quad (15)$$
Since \( \{x_k\} \) and \( \{d_k\} \) are bounded, so we can select a sub-sequence, let \( k \to \infty \) in (15), we obtain
\[
-\langle F(\hat{x}) , \hat{d} \rangle \leq 0
\]
(16)
Such that \( \hat{x} \) and \( \hat{d} \) are limits of subsequences that chosen. Otherwise, by (7) and already familiar argument,
\[
-\langle F(\hat{x}) , \hat{d} \rangle > 0
\]
(17)
(16) and (17) are a contradiction. Hence it is not possible to get that
\[
\lim \inf_{k \to \infty} \| F(x_k) \| > 0.
\]
This completes the proof.

4. Numerical Results
Here, we compared the performance of the new method (KA) previously discussed with the following two algorithms

SCGP: It's coming from [20].
SGPM: It's coming from [19].

All the codes are written in MATLAB with version R2014a, also all experiments have been run on a computer with 4 GB of RAM and CPU 2.30 GHz. To compare the results of the new algorithm (KA) with the algorithms mentioned above, we have run these codes.

This comparison is based on two things: number of iterations \( N_i \), number of functions evaluations \( N_f \). Also, the dimensions used to compare the algorithms are limited to 5000 - 50000 for the following initial points:

\[ x_0 = (10,10,\ldots,10)^T \]
\[ x_1 = (-10,-10,\ldots,-10)^T \]
\[ x_2 = (1,1,\ldots,1)^T \]
\[ x_3 = (-1,-1,\ldots,-1)^T \]
\[ x_4 = \left( \frac{1}{n}, \frac{1}{3}, \ldots, \frac{1}{n} \right)^T \]
\[ x_5 = (0.1,0.1,\ldots,0.1)^T \]
\[ x_6 = \left( \frac{1}{n}, \frac{2}{n}, \ldots, 1 \right)^T \]
\[ x_7 = (1 - \frac{1}{n}, 1 - \frac{2}{n}, \ldots, 0)^T. \]

When the total of \( N_i \) exceeds 500000 or \( \| F_k \| \leq 10^{-8} \) or \( \| F(z_k) \| \leq 10^{-8} \) then all the algorithms will be end. The parameters used for comparison were determined as follows \( \beta = 0.9 , \rho = 0.5 , \sigma = 0.7 , \epsilon = 10^{-8} , \mu = 0.2 , \theta = 1.5 , r = 0.2 \).

Digital experiments are displayed in the following table (4.1), Where it contains both \( N_i \) and \( N_f \) for all algorithms.

### Table 4.1. Numerical results \( (N_i, N_f) \)

| P. | Dim. | S.P | \( N_i \) | \( N_f \) | \( N_i \) | \( N_f \) | \( N_i \) | \( N_f \) | \( N_i \) | \( N_f \) |
|----|------|-----|--------|--------|--------|--------|--------|--------|--------|--------|
| 50000 | \( x_0 \) | 18 | 50 | 37 | 152 | 23 | 51 |
| 50000 | \( x_1 \) | 18 | 50 | 37 | 152 | 23 | 51 |
| 50000 | \( x_2 \) | 15 | 42 | 34 | 139 | 19 | 41 |
| 50000 | \( x_3 \) | 15 | 42 | 34 | 139 | 19 | 41 |
| 50000 | \( x_4 \) | 7 | 17 | 22 | 91 | 12 | 27 |
| 50000 | \( x_5 \) | 11 | 30 | 29 | 119 | 16 | 35 |
| 50000 | \( x_6 \) | 14 | 39 | 33 | 135 | 18 | 39 |
| 50000 | \( x_7 \) | 14 | 39 | 33 | 135 | 18 | 39 |
Table 4.1. Numerical results ($N_i, N_f$) – continued

| P   | Dim. | S.P | KA | SCGP | SGPM |
|-----|------|-----|----|------|------|
|     |      |     | $N_i$ | $N_f$ | $N_i$ | $N_f$ | $N_i$ | $N_f$ |
| $P_2$ | 50000 | $x_0$ | 18 | 50 | 37 | 152 | 24 | 51 |
|      | 50000 | $x_1$ | 13 | 36 | 23 | 117 | 16 | 36 |
|      | 50000 | $x_2$ | 15 | 42 | 34 | 139 | 20 | 41 |
|      | 50000 | $x_3$ | 11 | 31 | 21 | 109 | 13 | 28 |
|      | 50000 | $x_4$ | 7  | 17 | 22 | 91  | 13 | 27 |
|      | 50000 | $x_5$ | 11 | 30 | 29 | 119 | 17 | 35 |
|      | 50000 | $x_6$ | 14 | 39 | 33 | 135 | 19 | 39 |
|      | 50000 | $x_7$ | 14 | 39 | 33 | 135 | 19 | 39 |
| $P_3$ | 10000 | $x_0$ | 73648 | 814589 | 49213 | 262471 | 41422 | 287519 |
|      | 10000 | $x_1$ | 82356 | 912566 | 54607 | 291239 | 46013 | 320362 |
|      | 10000 | $x_2$ | 59460 | 656970 | 40198 | 214391 | 33813 | 235180 |
|      | 10000 | $x_3$ | 70904 | 784765 | 47452 | 253079 | 39967 | 277857 |
|      | 10000 | $x_4$ | 59952 | 662036 | 40778 | 209687 | 33070 | 230256 |
|      | 10000 | $x_5$ | 57756 | 637885 | 39316 | 206867 | 33070 | 230256 |
|      | 10000 | $x_6$ | 25485 | 281529 | 17281 | 92167 | 14545 | 101567 |
|      | 10000 | $x_7$ | 25586 | 282568 | 17278 | 92151 | 14465 | 99921 |
| $P_4$ | 10000 | $x_0$ | 14 | 41 | 20 | 104 | 15 | 37 |
|      | 10000 | $x_1$ | 17 | 50 | 45 | 220 | 28 | 85 |
|      | 10000 | $x_2$ | 11 | 31 | 18 | 94  | 15 | 38 |
|      | 10000 | $x_3$ | 16 | 50 | 33 | 163 | 26 | 83 |
|      | 10000 | $x_4$ | 12 | 34 | 29 | 145 | 23 | 73 |
|      | 10000 | $x_5$ | 12 | 37 | 33 | 163 | 31 | 107 |
|      | 10000 | $x_6$ | 13 | 39 | 27 | 135 | 22 | 66 |
|      | 10000 | $x_7$ | 14 | 43 | 27 | 135 | 22 | 66 |
| $P_5$ | 50000 | $x_0$ | 1482 | 4651 | 4724 | 11694 | 2442 | 6104 |
|      | 50000 | $x_1$ | 561 | 3233 | 1547 | 16348 | 927  | 3716 |
|      | 50000 | $x_2$ | 1566 | 4509 | 4196 | 9838  | 2352 | 5540 |
|      | 50000 | $x_3$ | 385 | 1878 | 707  | 2815  | 765  | 2514 |
|      | 50000 | $x_4$ | 1223 | 2499 | 2247 | 4516  | 4925 | 9853 |
|      | 50000 | $x_5$ | 932  | 2572 | 2288 | 4673  | 3778 | 7593 |
|      | 50000 | $x_6$ | 278  | 1251 | 524  | 2023  | 572  | 1790 |
|      | 50000 | $x_7$ | 1001 | 2263 | 4732 | 9717  | 3504 | 7032 |
| $P_6$ | 50000 | $x_0$ | 18  | 51  | 39  | 159  | 23  | 47  |
|      | 50000 | $x_1$ | 19  | 54  | 41  | 167  | 23  | 47  |
|      | 50000 | $x_2$ | 16  | 45  | 36  | 147  | 21  | 43  |
|      | 50000 | $x_3$ | 17  | 48  | 38  | 155  | 22  | 45  |
|      | 50000 | $x_4$ | 17  | 48  | 37  | 151  | 21  | 43  |
|      | 50000 | $x_5$ | 17  | 48  | 37  | 151  | 21  | 43  |
|      | 50000 | $x_6$ | 16  | 45  | 37  | 151  | 21  | 43  |
|      | 50000 | $x_7$ | 16  | 45  | 37  | 151  | 21  | 43  |
In addition to the results obtained in the table above, in order to obtain a comprehensive comparison, we will use the performance profile that was presented by More’ and Dolan [21] as a tool for evaluating the three algorithms and comparing them in terms of efficiency and durability.

![Figure 4.1. Performance profile for the total number of function evaluation](image)
Figure 4.2. Performance profile for the total number of iterations

Through the results of comparisons, the new method demonstrated its superiority to solve non-linear systems of monotonous equations compared with other methods used in comparison.

Figure (4.1) shows the performance for the total of $N_f$ for the three algorithms, fig.(4.2) shows the performance for the total of $N_i$. The algorithm (KA) solved about 59% and 86% of the test functions respectively, and it has least of $N_i$ and $N_f$ among the three methods and will reach to (1) faster than the other algorithms. This implies that the new algorithm (KA) is the best algorithm closing to the performance index.

5. Conclusion

Through the comparison technique and through the numerical results presented in the table above for problems with different points and initial distances, we conclude that the performance of the proposed algorithm (KA) more efficient and effective than the proposed algorithms for comparison in terms of $N_i$ and $N_f$.

In turn, it is possible to improve the behavior of the proposed new algorithm to solve nonlinear monotone equations that do not require information about Jacobian. Also, algorithm (KA) can find the best solution to problem (1) and has been created a global convergence without using any merit functions.

References

[1] A. N. Iusem and M. V. Solodov, "Newton-type methods with generalized distances for constrained optimization", Optimization, vol. 41, no. 3, pp. 257–278, 1997.

[2] Y. B. Zhao and D. Li, "Monotonicity of fixed point and normal mapping associated with variational inequality and its application", SIAM Journal on Optimization, vol. 11, no. 4, pp. 962–973, 2001.

[3] J. E. Dennis and J. J. Moré, "A characterization of superlinear convergence and its application to quasi-Newton methods", Mathematics of computation, vol. 28, no. 126, pp. 549–560, 1974.

[4] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables. New York: Academic Press, 1970.

[5] H. Jiang and D. Ralph, "Global and local superlinear convergence analysis of Newton-type
methods for semismooth equations with smooth least squares", Reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods. Springer, Boston, MA, pp. 181-209, 1998.

[6] K. Amini, M. A. Shiker and M. Kimiaei, "A line search trust-region algorithm with nonmonotone adaptive radius for a system of nonlinear equations", 4 OR- Journal of operation research, vol. 14, no. 2, pp. 133-152, 2016.

[7] A. Griewank, "The global convergence of Broyden-like methods with a suitable line search", The ANZIAM Journal, vol. 28, no. 1, pp. 75–92, 1986.

[8] M. V. Solodov and B. F. Svaiter, "A globally convergent inexact Newton method for systems of monotone equations", in: Fukushima M, Qi . (Eds.) Reformulation: Nonsmooth, Piecewise smooth, Semismooth and Smoothing Methods. Kluwer Academic Publishers, pp 355–369, 1998.

[9] D. H. Li and M. Fukushima, "A globally and superlinear convergent Gauss–Newton-based BFGS methods for symmetric nonlinear equations", SIAM Journal on Numerical Analysis, vol. 37, no. 1, pp. 152–172, 1999.

[10] M. A. Shiker and K. Amini, "A New Projection- Based Algorithm for Solving a Large-scale Nonlinear System of Monotone Equations", Croatian Operational Research Review, vol. 9, no. 1, pp. 63-73, 2018.

[11] J. Nocedal, "Updating quasi-Newton matrixes with limited storage", Mathematics of computation, vol. 35, no. 151, pp. 773–782, 1980.

[12] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization methods", Mathematical programming, vol. 45, no. (1-3), pp. 503–528, 1989.

[13] F. Facchinei and J. S Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems. NewYork: Springer, 2003.

[14] L. Qi, X. J. Tong and D. H. Li, "An active-set projected trust region algorithm for box constrained nonsmooth equations", Journal of Optimization Theory and Applications, vol. 120, no. 3, pp. 601–625, 2004.

[15] M. A. Shiker and Z. Sahib, "A modified Technique for Solving Unconstrained Optimization", Journal of Engineering and Applied Sciences, vol. 13, no. 22, pp. 9667-9671, 2018.

[16] J. Barzilai and J. M. Borwein, "Two point stepsize gradient methods", IMA journal of numerical analysis, vol. 8, no. 1, pp. 141–148, 1988.

[17] W. Cruz and M. Raydan, "Nonmonotone spectral methods for large-scale nonlinear systems", Optimization Methods and Software, vol. 18, no. 5, pp. 583–599, 2003.

[18] W. La Cruz, J. M. Martinez and M. Raydan, "Spectral residual method without gradient information for solving large-scale nonlinear systems of equations", Mathematics of Computation, vol. 75, no. 255, pp. 1429–1448, 2006.

[19] L. Zhang and W. J. Zhou, "Spectral gradient projection method for solving nonlinear monotone equations", Journal of Computational and Applied Mathematics, vol. 196, no. 2, pp. 478–484, 2006.

[20] M. Koorapetse, P. Kaelo and E. R. Offen, "A scaled derivative-free projection method for solving nonlinear monotone equations", Bulletin of the Iranian Mathematical Society, vol. 45, no. 3, pp. 755–770, 2019.

[21] E. D. Dolan and J. J. More', "Benchmarking optimization software with performance profiles", Mathematics of computation, vol. 91, no. 2, pp. 201–213, 2002.