Research Article

A Modified Scaled Spectral-Conjugate Gradient-Based Algorithm for Solving Monotone Operator Equations

Auwal Bala Abubakar 1,2, Kanikar Muangchoo 3, Abdulkarim Hassan Ibrahim 4, Sunday Emmanuel Fadugba 5, Kazeem Olalekan Aremu 2,6, and Lateef Olakunle Jolaoso 2

1Department of Mathematical Sciences, Faculty of Physical Sciences, Bayero University, Kano, Kano, Nigeria
2Department of Mathematics and Applied Mathematics, Sefako Makgatho Health Sciences University, Ga-Rankuwa, Pretoria 0204, South Africa
3Faculty of Science and Technology, Rajamangala University of Technology Phra Nakhon (RMUTP), 1381, Pracharat 1 Road, Wongawang, Bang Sue, Bangkok 10800, Thailand
4KMUTT Fixed Point Research Laboratory, Room SCL 802 Fixed Point Laboratory, Science Laboratory Building, Department of Mathematics, Faculty of Science, King Mongkut’s University of Technology Thonburi (KMUTT), 126 Pracha-Uthit Road, Bang Mod, Thung Khru, Bangkok 10140, Thailand
5Department of Mathematics, Ekiti State University, Ado Ekiti 360001, Nigeria
6Department of Mathematics, Faculty of Science, Usmanu Danfodiyo University Sokoto, Sokoto, Nigeria

Correspondence should be addressed to Kanikar Muangchoo; kanikar.m@rmutp.ac.th

Received 13 January 2021; Revised 11 April 2021; Accepted 13 April 2021; Published 26 April 2021

Academic Editor: Jen-Chih Yao

Copyright © 2021 Auwal Bala Abubakar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a modified scaled spectral-conjugate-based algorithm for finding solutions to monotone operator equations. The algorithm is a modification of the work of Li and Zheng in the sense that the uniformly monotone assumption on the operator is relaxed to just monotone. Furthermore, unlike the work of Li and Zheng, the search directions of the proposed algorithm are shown to be descent and bounded independent of the monotonicity assumption. Moreover, the global convergence is established under some appropriate assumptions. Finally, numerical examples on some test problems are provided to show the efficiency of the proposed algorithm compared to that of Li and Zheng.

1. Introduction

We desire in this work to propose an algorithm to solve the problem:

\[ F(x) = 0, \quad x \in \mathbb{C}, \]

(1)

where \( F: \mathbb{R}^n \rightarrow \mathbb{R}^n \) is monotone and Lipschitz continuous and \( \mathbb{C} \subseteq \mathbb{R}^n \) is nonempty, closed, and convex.

Solving problems of form (1) are becoming interesting in recent years due to its appearance in many areas of science, engineering, and economy, for example, in forecasting of financial market [1], constrained neural networks [2], economic and chemical equilibrium problems [3, 4], signal and image processing [5, 6], phase retrieval [7, 8], power flow equations [9], nonnegative matrix factorisation [10, 11], and many more.

Some notable methods for finding solution to (1) are: Newton’s method, quasi-Newton method, Gauss–Newton method, Levenberg–Marquardt method, and their variants [12–15]. These methods are prominent due to their fast convergence property. However, their convergence is local, and they require computing and storing of the Jacobian matrix at each iteration. In addition, there is a need to solve a linear equation at each iteration. These and other reasons make them unattractive especially for large-scale problems. To avoid the above drawbacks, methods that are globally convergent and also do not require computing and storing of the Jacobian matrix were
introduced. Examples of such methods are the spectral (SG) and conjugate (CG) gradient methods. However, SG and CG methods for solving (1) are usually combined with the projection method proposed in [16]. For instance, Zhang and Zhou [17] extended the work of Birgin and the projection method proposed in [16]. For instance, and CG methods for solving (1) are usually combined with (SG) and conjugate (CG) gradient methods. However, SG introduced. Examples of such methods are the spectral directions defined by Li and Zheng. In addition, the global monotone. q&_his is achieved by modifying the two search condition on the operator from uniformly monotone to on the operator. q&_he search directions defined in [24] were shown to be bounded under the uniformly monotone assumption. q&_he search directions defined in [24] were shown to be bounded under the uniformly monotone assumption. Our main interest is to modify the search directions defined in [24] and prove their boundedness without requiring the uniformly monotone assumption. The directions in [24] are defined as follows:

\[ d_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, 2, \ldots, \]

(2)

where \( x_{k+1} \) and \( x_k \) are the current and previous points, respectively. \( \alpha_k \) is the stepsize obtained via a line search and \( d_k \) is the search direction defined as

\[ d_0 = -F_0, \quad d_k = -\theta_k F_k + \beta_k d_{k-1} + \gamma_k y_{k-1}, \quad k \geq 1, \]

(3)

where \( \theta_k, \beta_k, \) and \( y_k \) are parameters and \( y_{k-1} = F_{k-1} - F_k \).

Based on the three-term direction above, we will propose a modified scaled three-term derivative-free algorithms for solving (1). The algorithms are a modification of the two algorithms proposed by Li and Zheng [24]. The aim of the modification is to relax the uniformly monotone assumption on the operator. The search directions defined in [24] were shown to be bounded under the uniformly monotone assumption. Our main interest is to modify the search directions defined in [24] and prove their boundedness without requiring the uniformly monotone assumption. The directions in [24] are defined as follows:

\[ d_k = -\mu_k F_k + \frac{1}{d_{k-1} y_{k-1}} \left( \mu_k F_k y_{k-1} - c_{k-1} \| y_{k-1} \|^2 \right) d_{k-1} + (2 - \mu_k) c_{k-1} y_{k-1}. \]

(4)

**STDF2:**

\[ d_k = -\mu_k F_k + \frac{1}{d_{k-1} y_{k-1}} \left( \mu_k F_k y_{k-1} - c_{k-1} \| y_{k-1} \|^2 \right) d_{k-1} + (2 - \mu_k) c_{k-1} y_{k-1}. \]

(5)
where
\[
\begin{align*}
\alpha_{k-1} &= \frac{F^T_k d_{k-1}}{d^T_{k-1} y_{k-1}}, \\
\beta_{k-1} &= \frac{F^T_k d_{k-1}}{d^T_{k-1} y_{k-1}}, \\
y_{k-1} &= F_k y_{k-1}, \\
\xi_{k-1} &= y_{k-1} + \eta_{k-1} d_{k-1}, \\
\rho_{k-1} &= 1 + \max \left\{ 0, -\frac{d^T_{k-1} y_{k-1}}{\|d_{k-1}\|^2} \right\}.
\end{align*}
\]

Remark 1.
\[
d^T_{k-1} y_{k-1} \geq d^T_{k-1} y_{k-1} + \|d_{k-1}\|^2 - d^T_{k-1} y_{k-1} = \|d_{k-1}\|^2.
\]

From (10), a lower bound for the term \(d^T_{k-1} y_{k-1}\) is obtained without any assumption on the operator \(F\).

Let \(\operatorname{Sol}(\mathcal{C}, F)\) be the solution set of (1) and assume that the following holds.

**Assumption 1.** The constraint set \(\mathcal{C}\) is nonempty, closed, and convex.

**Assumption 2.** The operator \(F\) is monotone, that is, \(\forall x_1, x_2 \in \mathbb{R}^n:\)
\[
(F(x_1) - F(x_2))^T, \quad (x_1 - x_2) \geq 0.
\]

To obtain a lower bound for the term \(d^T_{k-1} y_{k-1}\), Li and Zheng used the uniformly monotone assumption. So, in order to relax this condition, we replace the term \(d^T_{k-1} y_{k-1}\) in the directions defined by (4) and (5) with \(d^T_{k-1} w_{k-1}\). In addition, we replace \(c_{k-1}\) and \(\tau\) in (4) and (5) with \(\beta_{k-1}, \xi_{k-1}\) in (5) with \(d_{k-1}\). Hence, we define the new directions as follows:

**PSTDF1:**
\[
d_k := -\mu_k F_k + \frac{1}{d^T_{k-1} w_{k-1}} \left( \mu_k F_k^T y_{k-1} - \beta_{k-1} \|y_{k-1}\|^2 \right) d_{k-1} + (2 - \mu_k) \beta_{k-1} y_{k-1}.
\]

**PSTDF2:**
\[
d_k := -\mu_k F_k + \frac{1}{d^T_{k-1} w_{k-1}} \left( \mu_k F_k^T y_{k-1} - \beta_{k-1} \|y_{k-1}\|^2 - \mu_2 F_k^T d_{k-1} \right) d_{k-1} + (2 - \mu_k) \beta_{k-1} y_{k-1}.
\]

Assumption 3. The operator \(F\) is L-Lipschitz continuous on \(\mathbb{R}^n\), that is, \(\forall x_1, x_2 \in \mathbb{R}^n, L > 0,\)
\[
\|F(y_1) - F(y_2)\| \leq L \|x_1 - x_2\|.\]

In the following algorithm, we generate approximate solutions to problem (1) under Assumptions 1–3.

**Algorithm 1.** PSTDF.

**Input.** Choose an initial guess \(x_0 \in \mathcal{C}, \theta > 0, 0 < \rho < 1, 1 < \mu_1 \leq 2, \mu_2 \geq 0, t > 0, \text{tol} > 0\) and \(k = 0\).

**Step 1.** If \(\|F(x_k)\| \leq \text{tol}, \) terminate. Else move to **Step 2.**

**Step 2.** Compute \(d_k\) using (7) or (8).

**Step 3.** Compute
\[
z_k = x_k + \alpha_k d_k,
\]
\[\alpha_k = \theta \rho^i, \text{ for } i = 0, 1, \ldots, \text{ where } i \text{ is the least nonnegative integer satisfying}
\]
\[
-F(z_k)^T d_k \geq t \alpha_k \|d_k\|^2.
\]

**Step 4.** If \(z_k \in \mathcal{C}\) and \(\|F(z_k)\| \leq \text{tol}, \) then stop. Else, compute
\[ x_{k+1} := P_{\mathcal{E}}[x_k - \eta_k F(z_k)], \quad (15) \]

where
\[ \eta_k := \frac{F(z_k)^T (x_k - z_k)}{\|F(z_k)\|^2}. \tag{16} \]

**Step 5.** Let \( k = k + 1 \) and repeat from **Step 1.**

3. **Theoretical Results**

In this section, we will establish the convergence analysis of the proposed algorithm. However, we require the following important lemmas. The following lemma shows that the proposed directions are descent.

**Lemma 1.** The search directions defined by (7) and (8) satisfy the sufficient descent condition.

*Proof.* Multiplying both sides of (7) by \( F_k^T \), we have

\[
F_k^T d_k = -\mu_k \|F_k\|^2 + \frac{1}{d_{k-1}^T w_{k-1}} \left( \mu_k F_k^T y_{k-1} - c_k \|y_{k-1}\| \right)^2 F_k^T d_{k-1} + (2 - \mu_k) c_k F_k^T y_{k-1} F_k^T d_{k-1} \leq -\mu_k \|F_k\|^2.
\]

Also, multiplying both sides of (8) by \( F_k^T \), we have

\[
F_k^T d_k = -\mu_k \|F_k\|^2 + \frac{1}{d_{k-1}^T w_{k-1}} \left( \mu_k F_k^T y_{k-1} - c_k \|y_{k-1}\| \right)^2 F_k^T d_{k-1} + (2 - \mu_k) c_k F_k^T y_{k-1} F_k^T d_{k-1} \leq -\mu_k \|F_k\|^2.
\]

Hence, for all \( k \), the directions defined by (7) and (8) satisfy

\[ F_k^T d_k \leq - (\mu_k - 1) \|F_k\|^2. \tag{19} \]

The lemma below shows that the linesearch (14) is well-defined and the stepsize is bounded away from zero.

**Lemma 2** (see [5]). Suppose Assumptions 1–3 are satisfied. If \( \{d_k\}, \{z_k\}, \) and \( \{x_k\} \) are sequences defined by (7), (13), and (15), respectively, then

(i) For all \( k \), there is \( \alpha_k = \theta \rho^k \) satisfying (14) for some \( \theta \in \mathbb{N} \cup \{0\} \) and \( \forall k \geq 0 \).

(ii) \( \alpha_k \) obtained via (14) satisfies
\[ \alpha_k > \max \left\{ 0, \frac{\rho \| F_k \|^2}{(L + \rho) \| d_k \|^2} \right\}. \]  

(20)

**Lemma 3** (see [5]). Suppose Assumptions 1–3 are fulfilled, then the sequences \( \{z_k\} \) and \( \{x_k\} \) defined by (13) and (15) are bounded. Furthermore, \[ \lim_{k \to \infty} \| x_k - z_k \| = \lim_{k \to \infty} \| w_k \| d_k \| = 0. \]  

(21)

**Lemma 4** (see [5]). From Lemma 3, we have \[ \| x_{k+1} - x \|^2 < \| x_k - x \|^2. \]  

(22)

**Remark 2.** Since \( \{x_k\}\) is bounded from Lemma 3 and \( F \) is continuous from Assumption 3, \( \| F_k \| \) is also bounded. That is, there exists \( c_1, c_2 > 0 \) such that, for all \( k \), \[ \| x_k \| \leq c_1, \| F_k \| \leq c_2. \]  

(23)

All are now set to establish the convergence of the proposed algorithm.

\[ \| d_k \| \leq \mu_1 \| F_k \| + \frac{\| d_{k-1} \|}{d_{k-1}^T w_{k-1}} \left( \mu_1 \| F_k \| \| y_{k-1} \| + \frac{\| F_k \| \| d_{k-1} \| \| y_{k-1} \|}{d_{k-1}^T w_{k-1}} \right) + 2 \frac{\| F_k \| \| d_{k-1} \| \| y_{k-1} \|}{d_{k-1}^T w_{k-1}} + \mu_1 \frac{\| F_k \| \| d_{k-1} \| \| y_{k-1} \|}{d_{k-1}^T w_{k-1}} \]  

\[ \leq \mu_1 \| F_k \| + 2 (\mu_1 + 1) \frac{L \| x_k - x_{k-1} \|}{d_{k-1}^T w_{k-1}} + \left( \frac{L \| x_k - x_{k-1} \|}{d_{k-1}^T w_{k-1}} \right)^2 \| F_k \| \]  

\[ \leq \mu_1 c_2 + 2 (\mu_1 + 1) \frac{L (2c_1)}{(\mu_1 - 1) \nu} + \left( \frac{L (2c_1)}{(\mu_1 - 1) \nu} \right)^2 c_2 \]  

\[ \leq \mu_1 c_2 + 2 (\mu_1 + 1) \frac{L (2c_1)}{(\mu_1 - 1) \nu} \]  

(24)

**Theorem 1.** Suppose Assumptions 1 and 2 are satisfied. If \( \{x_k\} \) is a sequence defined by (15), then \[ \lim \inf_{k \to \infty} \| F_k \| = 0. \]  

(25)

Furthermore, the sequence \( \{x_k\}\) converges to a solution of problem (1).

**Proof.** Suppose that \( \lim \inf_{k \to \infty} \| F_k \| \neq 0 \), then there is a positive constant \( \nu > 0 \) such that, for all \( k \geq 0 \), \[ \| F_k \| \geq \nu. \]  

(26)

By (17), (18), and the Cauchy–Schwartz inequality, we have that, for all \( k \geq 0 \), \[ \| d_k \| \geq (\mu_1 - 1) \| F_k \| \geq (\mu_1 - 1) \nu. \]  

(27)

To complete the proof of the theorem, we need to show that the search direction \( d_k \) defined by (7) and (8) are bounded.

For \( k = 0 \), we have \[ \| d_0 \| = \| F_0 \| \leq c_2. \]  

(28)

Now for \( k \geq 1 \), using (7), (10), (12), and (26), we have

\[ \| d_k \| \leq (\mu_1 + \mu_2) \| F_k \| + 2 (\mu_1 + 1) \frac{L \| x_k - x_{k-1} \|}{d_{k-1}^T w_{k-1}} + \left( \frac{L \| x_k - x_{k-1} \|}{d_{k-1}^T w_{k-1}} \right)^2 \| F_k \| \]  

\[ \leq (\mu_1 + \mu_2) c_2 + 2 (\mu_1 + 1) \frac{L (2c_1)}{(\mu_1 - 1) \nu} \]  

(29)
Letting
\[ M_1 = \mu_1 c_2 + (4c_1 L (\mu_1 + 1)/(\mu_1 - 1)^2) c_2 \] and
\[ M_2 = (\mu_1 + \mu_2) c_2 + (4c_1 L (\mu_1 + 1)/(\mu_1 - 1)^2) + (4c_2 L^2)/(\mu_1 - 1)^2 c_2, \] then for all \( k \),
\[ \frac{d_k}{d_k} \leq M_2, \quad (30) \]

\[ d_k \leq M_2, \quad (30) \]

since \( M_2 > M_1 \).

Multiplying (20) by \( \|d_k\| \), we get
\[ \alpha_k \|d_k\| \geq \max \left\{ 0 \|d_k\|, \frac{\rho \|F_k\|^2}{(L + t)\|d_k\|} \right\} \geq \max \left\{ \left( \mu_1 - 1 \right)^2 \frac{\mu (\mu_1 - 1)^2 \nu^2}{(L + t)M} \right\} > 0. \quad (31) \]

This contradicts (21) and hence \( \lim_{k \to \infty} \inf \|F_k\| = 0. \)

Because \( F \) is a continuous function and (24) holds, then the sequence \( \{x_k\} \) has some accumulation point say \( x \) for which \( F(x) = 0 \), that is, \( x \) is a solution of (1). From (22), it holds that \( \|x_k - x\| \) converges, and since \( x \) is an accumulation point of \( \{x_k\} \), \( \{x_k\} \) converges to \( x \).  

4. Numerical Examples on Monotone Operator Equations

This segment of the paper would demonstrate the computational efficiency of the PSTDF algorithm relative to STDF algorithm [24]. For PSTDF algorithm, we have PSTDF1 which corresponds to the direction defined by (7) and PSTDF2 corresponding to the one defined by (8). Similarly, for the STDF algorithm, we have STDF1 and STDF2 corresponding to (4) and (5), respectively. The parameters chosen for the implementation of the PSTDF algorithm are \( \theta = 1, \mu_1 = 1.9, \mu_2 = 0.8, \rho = 0.8, \) and \( t = 10^{-4} \). The parameters for STDF algorithm are chosen as reported in [24]. The metrics considered are the number of iteration (NOI), number of function evaluations (NFE), and the CPU time (TIME). We used eight test problems with dimension \( n = 1000, 5000, 10,000, 50,000, \) and \( 100,000 \) and five initial points \( x_1 = (0.1, 0.1, \ldots, 0.1)^T, \quad x_2 = (0.2, 0.2, \ldots, 0.2)^T, \quad x_3 = (0.5, 0.5, \ldots, 0.5)^T, \quad x_4 = (1.5, 1.5, \ldots, 1.5)^T, \quad x_5 = (2, 2,
The algorithms were coded in MATLAB R2019a and run on a PC with Intel (R) Core (TM) i3-7100U processor with 8 GB RAM and CPU 2.40 GHz. The iteration process is stopped whenever $\|F(x_k)\| \leq 10^{-5}$. Failure is declared if this condition is not satisfied after 1000 iterations.

Table 1 consists of the test problems considered, where the function $F$ is $F(x) = (f_1(x), f_2(x), \ldots, f_n(x))^T$ and $x = (x_1, x_2, \ldots, x_n)^T$.

The result of the experiments in Tabular form can be found in the link https://documentcloud.adobe.com/link/review?uri=urn:aaid:scds:US:77a9a900-2156-4344-a9d9-b42e3a3dc8e5. It can be observed from the results that the algorithms successfully solved all the problems considered without a single failure. However, to better illustrate the performance of each algorithm, we employ the Dolan and Moré [47] performance profiles and plot Figures 1–3. Figures 1–3 represent the performance of the algorithms based on NOI, NFE, and TIME, respectively. In terms of NOI (Figure 1), the best performing algorithm is PSTDF2 with 70% success, followed by PSTDF1 with 51% success. STDF1 and STDF2 record less than 10% success each. Based on NFE (Figure 2), the best performing algorithm is PSTDF1 with around 42% success, followed by PSTDF2 with almost 40% success. STDF1 and STDF2 record less than 10% success each. In terms of TIME (Figure 3), PSTDF2 performs better with around 50% success, followed by PSTDF1 with more than 30% success. STDF1 and STDF2 record around 20% and 5% success, respectively.
Overall, we can conclude that PSTDF1 and PSTDF2 outperform STDF1 and STDF2 based on the metrics considered.

5. Conclusions

In this paper, a modified scaled algorithm based on the spectral-conjugate gradient method for solving nonlinear monotone operator equations was proposed. The algorithm replaces the stronger assumption of uniformly monotone on the operator in the work of Li and Zheng (2020) with just monotone, which is weaker. Interestingly, the search directions were shown to be descent independent of line search and also without monotonicity assumption (unlike in the work of Li and Zheng). Furthermore, the convergence results were established under monotonicity and Lipschitz continuity assumptions on the operator. Numerical experiments on some benchmark problems were conducted to illustrate the good performance of the proposed algorithm.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

All authors contributed equally to the manuscript and read and approved the final manuscript.

Acknowledgments

The first, fifth, and the sixth authors acknowledge with thanks, the Department of Mathematics and Applied Mathematics at the Sefako Makgatho Health Sciences University. The second author was financially supported by the Rajamangala University of Technology Phra Nakhon (RMUTP) Research Scholarship.

References


