Applied Mathematics and Computation 294 (2017) 169–179

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Convergence rate results for steepest descent type method for nonlinear ill-posed equations

Santhosh George*, M. Sabari

Department of Mathematical and Computational Sciences, NIT Karnataka, Mangaluru 575 025, India

ARTICLE INFO

MSC:
65J15
65J20
47H17

Keywords:
Nonlinear ill-posed problem
Steepest descent method
Minimal error method
Regularization method
Discrepancy principle

ABSTRACT

Convergence rate result for a modified steepest descent method and a modified minimal error method for the solution of nonlinear ill-posed operator equation have been proved with noisy data. To our knowledge, convergence rate result for the steepest descent method and minimal error method with noisy data are not known. We provide a convergence rate results for these methods with noisy data. The result in this paper are obtained under less computational cost when compared to the steepest descent method and minimal error method. We present an academic example which satisfies the assumptions of this paper.

1. Introduction

Steepest descent method is used extensively (see [1–8] for linear ill-posed equations and [11,13] for nonlinear ill-posed equations) for solving ill-posed operator equations. In this study we consider a modified steepest descent method and a modified minimal error method for approximately solving the operator equation

$$F(x) = y, \tag{1.1}$$

where F: $D(F) \subseteq X \to Y$ is a nonlinear Fréchet differentiable operator between the Hilbert spaces X and Y. Let $D(F)$, $\langle ., .\rangle$ and $\|.\|$, respectively stand for the domain of F, inner product and norm which can always be identified from the context in which they appear. Fréchet derivative of F is denoted by $F'(.)$ and its adjoint by $F'(.)^*$. Further we assume that Eq. (1.1) has a solution $\hat{x}$, which is not depending continuously on the right-hand side data y, i.e., (1.1) is ill-posed.

It is assumed further that we have only approximate data $y^\delta \in Y$ with

$$\|y - y^\delta\| \le \delta.$$

Steepest descent method was considered by Scherzer [13], Neubauer and Scherzer [11] for approximately solving (1.1). In general, steepest-descent method for (1.1) can be written as

$$x_{k+1} = x_k + \alpha_k s_k, \tag{1.2}$$

where s_k is the search direction taken as the negative gradient of the minimization functional involved and α_k is the descent. For solving Eq. (1.1) with y^δ in place of y, method (1.2) was studied by Scherzer [13] when $s_k = -F'(x_k)^*(F(x_k) - y^\delta)$ and $\alpha_k = \frac{\|s_k\|^2}{\|F'(x_k)^* s_k\|^2}$. For linear operator F, Gilyazov [10] studied (α-process) method (1.2) when $s_k = -F'(x_k)^*(F(x_k) - y^\delta)$

* Corresponding author.
E-mail addresses: nitksanthosh@gmail.com, sgeorge@nitk.ac.in (S. George), msabarijothi16@gmail.com (M. Sabari).

http://dx.doi.org/10.1016/j.amc.2016.09.009
0096-3003/

and $\alpha_k = \frac{\langle (F^*F)^{\alpha} s_k, s_k\rangle}{\langle (F^*F)^{\alpha} s_k, F^*F s_k\rangle}$. Vasin [14] considered a regularized version of the steepest descent method in which $s_k = -F'(x_k)^*(F(x_k) - y^{\delta}) + \alpha(x_k - x_0)$ and $\alpha_k = \frac{\|s_k\|^2}{\|F'(x_k)s_k\|^2 + \alpha\|s_k\|^2}$. Here and below x_0 is the initial guess. Also, observe that the TIGRA-method of Ramlau [12] is of the form (1.2) with $s_k = -[F'(x_k)^*(F(x_k) - y^{\delta}) + \alpha_k(x_0 - x_k)]$ and $\alpha_k = \beta_k$. Note that, in all these methods, one has to compute Fréchet derivative of F at each iterate x_k in each iteration step which is in general very expensive.

1.1. Preliminaries

Let $B(x, r)$, $\overline{B}(x, r)$ stand, respectively for the open and closed balls in X , with center $x \in X$ and of radius $r > 0$.
In [11], Neubauer and Scherzer considered the steepest descent method:

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k s_k \quad (k = 0, 1, 2, \dots) \\ s_k &= -F'(x_k)^*(F(x_k) - y) \\ \alpha_k &= \frac{\|s_k\|^2}{\|F'(x_k)s_k\|^2} \end{aligned} \tag{1.3}$$

and the minimal error method:

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k s_k \quad (k = 0, 1, 2, \dots) \\ s_k &= -F'(x_k)^*(F(x_k) - y) \\ \alpha_k &= \frac{\|F(x_k) - y\|^2}{\|s_k\|^2} \end{aligned} \tag{1.4}$$

in the noise free case and obtained the rate

$$\|x_k - \hat{x}\| = O(k^{-\frac{1}{2}}) \tag{1.5}$$

under the assumptions $(\mathcal{A})$:

$(\mathcal{A}_1)$ F has a Lipschitz continuous Fréchet derivative $F'(.)$ in a neighborhood of x_0.
$(\mathcal{A}_2)$ $F'(x) = R_x F'(\hat{x}),\ x \in B(x_0, \rho)$ where $\{R_x\colon x \in B(x_0, \rho)\}$ is a family of bounded linear operators $R_x : Y \longrightarrow Y$ with

$$\|R_x - I\| \le C\|x - \hat{x}\|$$

where C is a positive constant and
$(\mathcal{A}_3)$

$$x_0 - \hat{x} = (F'(\hat{x})^* F'(\hat{x}))^{\frac{1}{2}} z$$

for some $z \in X$.

In the present paper, we consider a modified steepest descent method and a modified minimal error method, in the case of noise free case, defined by

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k s_k \quad (k = 0, 1, 2, \dots) \\ s_k &= -F'(x_0)^*(F(x_k) - y) \\ \alpha_k &= \frac{\|s_k\|^2}{\|F'(x_0)s_k\|^2} \end{aligned} \tag{1.6}$$

and

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k s_k \quad (k = 0, 1, 2, \dots) \\ s_k &= -F'(x_0)^*(F(x_k) - y) \\ \alpha_k &= \frac{\|F(x_k) - y\|^2}{\|s_k\|^2}, \end{aligned} \tag{1.7}$$

respectively. Instead of assumptions $(\mathcal{A})$, we use the following assumptions $(\mathcal{C})$:

$(\mathcal{C}_0)$ $\|F'(x)\| \le m$ for some $m > 0$ and for all $x \in D(F)$.
$(\mathcal{C}_1)$ $F'(\hat{x}) = F'(x_0)G(\hat{x}, x_0)$ where $G(\hat{x}, x_0)$ is a bounded linear operator from $X \longrightarrow X$ with

$$\|G(\hat{x}, x_0) - I\| \le C_0\rho$$

where C_0 is a positive constant and $\rho \ge \|x_0 - \hat{x}\|$.
$(\mathcal{C}_2)$ $F'(x) = R(x, y)F'(y)$ $(x, y \in B(x_0, \rho))$ where $\{R(x, y)\colon x, y \in B(x_0, \rho)\}$ is a family of bounded linear operators $R(x, y) : Y \longrightarrow Y$ with

$$\|R(x, y) - I\| \le C_1\|x - y\|$$

for some positive constant C_1 and

$(\mathcal{C}_3)$

$$x_0 - \hat{x} = (F'(x_0)^* F'(x_0))^{\frac{1}{2}} \nu$$

for some $\nu \in X$.

Observe that $x_0 - \hat{x}$ in $(\mathcal{C}_3)$ is depending on the known initial guess x_0, where as in $(\mathcal{A}_3)$, $x_0 - \hat{x}$ is depending on the unknown $\hat{x}$. Not only this advantage but also in our method one need to compute the Fréchet derivative only at one point x_0 throughout the iteration process.

As already mentioned in the abstract, no convergence rate results are known for steepest descent method and minimal error method with noisy data. In other words, it remains an open question whether convergence rate results can be proven for the methods (1.3) and (1.4) with noisy data. We considered methods (1.6) and (1.7) with noisy data and obtained a convergence rate result. Using the same idea we obtained a convergence rate result for methods (1.3) and (1.4).

The rest of the paper is organized as follows. In Section 2 we present a Convergence analysis of method (1.6) and (1.7) and in Section 3 we present a Convergence rate result for method (1.6) and (1.7) with noisy data. In particular, we consider a discrepancy principle in Section 3.1 and in Section 3.2 we present a convergence rate result for steepest descent method and minimal error method. Finally the paper ends with an academic example in Section 4.

2. Convergence analysis of method (1.6) and (1.7)

The main purpose of this section is to obtain an error estimate for $\|x_k - \hat{x}\|$ under the assumptions $(\mathcal{C})$. For this purpose we make use of the following result in [10] (see [10, Lemma 2]). Let $\{\upsilon_k\}$ be a sequence in X, $\nu > 0$, be some parameter such that

$$\|A^\nu \upsilon_k\|^2 - \|A^\nu \upsilon_{k+1}\|^2 \geq \varepsilon_k \langle A^{\nu+1}\upsilon_k, A^\nu \upsilon_k\rangle$$

for $k = 0, 1, 2, \ldots,$ where A is a positive self adjoint operator and $\varepsilon_k > 0$. Then

$$\|A^\nu \upsilon_k\| \leq [2(\nu+1)]^\nu \|\upsilon_k\|^{\frac{1}{\nu+1}} \left[\sum_{i=0}^{k-1} \varepsilon_i \|\upsilon_i\|^{-\frac{1}{\nu+1}}\right]^{-\nu}. \tag{2.1}$$

We shall apply the above result (i.e., (2.1)) to $\upsilon_k = A^{-\frac{1}{2}}(x_k - \hat{x})$ with $A = F'(x_0)^* F'(x_0)$ and $\nu = \frac{1}{2}$. Therefore, in order to apply (2.1), we need to prove;

$$\|x_k - \hat{x}\|^2 - \|x_{k+1} - \hat{x}\|^2 \geq \varepsilon_k \langle A(x_k - \hat{x}), x_k - \hat{x}\rangle \tag{2.2}$$

for some $\varepsilon_k > 0$ and $\|A^{-\frac{1}{2}}(x_k - \hat{x})\|$ is bounded. Let $\bar{C} = \max\{C_0, C_1\}$.

Lemma 2.1. *Let $(\mathcal{C})$ conditions hold and let $\bar{C}\rho \leq \sqrt{5} - 2$. Let x_k be as in (1.6) or (1.7). Then, $x_k \in B(x_0, \rho)$ and*

$$\|x_{k+1} - \hat{x}\|^2 + \alpha_k[1 - \bar{C}^2\rho^2 - 4\bar{C}\rho]\|A^{\frac{1}{2}}(x_k - \hat{x})\|^2 \leq \|x_k - \hat{x}\|^2 \tag{2.3}$$

for all $k = 0, 1, 2, \ldots$. Moreover,

$$\sum_{k=0}^{\infty} \alpha_k \|A^{\frac{1}{2}}(x_k - \hat{x})\|^2 < \infty.$$

Proof. We shall prove the result using induction. Note that $x_0 \in B(x_0, \rho)$ and suppose $x_k \in B(x_0, \rho)$. Then using (1.6) or (1.7), we have

$$\begin{aligned}
\|x_{k+1} - \hat{x}\|^2 - \|x_k - \hat{x}\|^2 &= -2\alpha_k \langle x_k - \hat{x}, F'(x_0)^*(F(x_k) - y)\rangle + \alpha_k^2 \|F'(x_0)^*(F(x_k) - y)\|^2 \\
&= -2\alpha_k \langle x_k - \hat{x}, F'(x_0)^*[F(x_k) - F(\hat{x}) - F'(x_0)(x_k - \hat{x})]\rangle \\
&\quad + \alpha_k[\alpha_k \|F'(x_0)^*(F(x_k) - y)\|^2 - 2\langle x_k - \hat{x}, F'(x_0)^* F'(x_0)(x_k - \hat{x})\rangle] \\
&= -2\alpha_k \left\langle F'(x_0)(x_k - \hat{x}), \int_0^1 (F'(\hat{x} + \theta(x_k - \hat{x})) - F'(x_0)) d\theta (x_k - \hat{x})\right\rangle \\
&\quad + \alpha_k[\alpha_k \|F'(x_0)^*(F(x_k) - y)\|^2 - 2\|A^{\frac{1}{2}}(x_k - \hat{x})\|^2].
\end{aligned} \tag{2.4}$$

So by $(\mathcal{C}_2)$, we have

$$\begin{aligned}
\|x_{k+1} - \hat{x}\|^2 - \|x_k - \hat{x}\|^2 &= -2\alpha_k \left\langle F'(x_0)(x_k - \hat{x}), \int_0^1 [R(\hat{x} + \theta(x_k - \hat{x}), x_0) - I] d\theta F'(x_0)(x_k - \hat{x})\right\rangle \\
&\quad + \alpha_k[\alpha_k \|F'(x_0)^*(F(x_k) - y)\|^2 - 2\|A^{\frac{1}{2}}(x_k - \hat{x})\|^2] \\
&\leq 2\alpha_k \int_0^1 \|R(\hat{x} + \theta(x_k - \hat{x}), x_0) - I\| \|F'(x_0)(x_k - \hat{x})\|^2 d\theta
\end{aligned}$$

$$
\begin{aligned}
&+ \alpha_k[\alpha_k\|F'(x_0)^*(F(x_k)-y)\|^2 - 2\|A^{\frac{1}{2}}(x_k-\hat{x})\|^2] \\
&\leq 2\alpha_k\bar{C}\|\hat{x}+\theta(x_k-\hat{x})-x_0\|\|A^{\frac{1}{2}}(x_k-\hat{x})\|^2 \\
&+ \alpha_k[\alpha_k\|F'(x_0)^*(F(x_k)-y)\|^2 - 2\|A^{\frac{1}{2}}(x_k-\hat{x})\|^2]. \qquad (2.5)
\end{aligned}
$$

Observe that $\alpha_k\|F'(x_0)^*(F(x_k)-y)\|^2 = \|F(x_k)-y\|^2$ in the case of method (1.7) and in the case of (1.6), we have $\alpha_k\|F'(x_0)^*(F(x_k)-y)\|^2 = \frac{\langle F'(x_0)s_k, F(x_k)-y\rangle^2}{\|F'(x_0)s_k\|^2} \leq \|F(x_k)-y\|^2$. So for both methods (1.6) and (1.7), we have

$$
\begin{aligned}
\alpha_k\|F'(x_0)^*(F(x_k)-y)\|^2 &\leq \|F(x_k)-y\|^2 \\
&= \left\|\int_0^1 F'(\hat{x}+\theta(x_k-\hat{x}))d\theta(x_k-\hat{x})\right\|^2 \\
&= \left\|\int_0^1 [R(\hat{x}+\theta(x_k-\hat{x}),x_0)-I+I]d\theta F'(x_0)(x_k-\hat{x})\right\|^2 \\
&\leq (\bar{C}\|\hat{x}+\theta(x_k-\hat{x})-x_0\|+1)^2\|F'(x_0)(x_k-\hat{x})\|^2 \\
&\leq (\bar{C}\rho+1)^2\|A^{\frac{1}{2}}(x_k-\hat{x})\|^2. \qquad (2.6)
\end{aligned}
$$

Therefore, by (2.5) and (2.6) we have

$$\|x_{k+1}-\hat{x}\|^2 - \|x_k-\hat{x}\|^2 \leq [(\bar{C}\rho+1)^2+2\bar{C}\rho-2]\alpha_k\|A^{\frac{1}{2}}(x_k-\hat{x})\|^2.$$

This completes the proof. □

Let

$$p(t) = -2t^3 - \frac{26}{5}t^2 + \frac{56}{5}t - \frac{4}{5}.$$

Note that $p(0) = -\frac{4}{5} < 0$ and $p(1) = \frac{16}{5} > 0$. So $p(t)$ has a zero in (0, 1). Let r_0 be the smallest zero of p in (0, 1). Next, we shall prove the boundedness of $\|A^{-\frac{1}{2}}(x_k-\hat{x})\|$.

Lemma 2.2. *Let* $(\mathcal{C})$ *conditions hold and* $\bar{C}\rho < \min\{\sqrt{5}-2, r_0\} = 0.0740$. *Let* x_k *be as in (1.6) or (1.7). Then,* $\|A^{-\frac{1}{2}}(x_k-\hat{x})\|$ *is bounded.*

Proof. Using $(\mathcal{C}_3)$, one can prove that $x_k-\hat{x}\in R(A^{\frac{1}{2}})$ for all $k=0,1,2,\ldots$. Therefore, we can apply the operator $A^{-\frac{1}{2}}$ to $x_{k+1}-\hat{x}$ and $x_k-\hat{x}$ to obtain

$$
\begin{aligned}
&\|A^{-\frac{1}{2}}(x_{k+1}-\hat{x})\|^2 - \|A^{-\frac{1}{2}}(x_k-\hat{x})\|^2 \\
&= -2\alpha_k\langle A^{-\frac{1}{2}}(x_k-\hat{x}), A^{-\frac{1}{2}}F'(x_0)^*(F(x_k)-y)\rangle + \alpha_k^2\|A^{-\frac{1}{2}}F'(x_0)^*(F(x_k)-y)\|^2 \\
&= -2\alpha_k\langle A^{-\frac{1}{2}}(x_k-\hat{x}), A^{-\frac{1}{2}}F'(x_0)^*(F(x_k)-F(\hat{x})-F'(\hat{x})(x_k-\hat{x}))\rangle \\
&\quad + \alpha_k[\alpha_k\|A^{-\frac{1}{2}}F'(x_0)^*(F(x_k)-y)\|^2 - 2\langle A^{-\frac{1}{2}}(x_k-\hat{x}), A^{-\frac{1}{2}}F'(x_0)^*F'(\hat{x})(x_k-\hat{x})\rangle] \\
&= -2\alpha_k\left\langle A^{-\frac{1}{2}}(x_k-\hat{x}), \int_0^1 (F'(\hat{x}+\theta(x_k-\hat{x}))-F'(\hat{x}))d\theta(x_k-\hat{x})\right\rangle \\
&\quad + \alpha_k[\alpha_k\|F(x_k)-y\|^2 - 2\langle A^{-\frac{1}{2}}(x_k-\hat{x}), F'(\hat{x})(x_k-\hat{x})\rangle]. \qquad (2.7)
\end{aligned}
$$

So by $(\mathcal{C}_2)$ and (2.7), we have

$$
\begin{aligned}
&\|A^{-\frac{1}{2}}(x_{k+1}-\hat{x})\|^2 - \|A^{-\frac{1}{2}}(x_k-\hat{x})\|^2 \\
&= -2\alpha_k\left\langle A^{-\frac{1}{2}}(x_k-\hat{x}), \int_0^1 (R(\hat{x}+\theta(x_k-\hat{x}),\hat{x})-I)d\theta F'(\hat{x})(x_k-\hat{x})\right\rangle \\
&\quad + \alpha_k[\alpha_k\|F(x_k)-y\|^2 - 2\langle A^{-\frac{1}{2}}(x_k-\hat{x}), F'(\hat{x})(x_k-\hat{x})\rangle]. \\
&= -2\alpha_k\left\langle A^{-\frac{1}{2}}(x_k-\hat{x}), \int_0^1 [R(\hat{x}+\theta(x_k-\hat{x}),\hat{x})-I]d\theta F'(\hat{x})(x_k-\hat{x})\right\rangle \\
&\quad + \alpha_k[\alpha_k\|F(x_k)-y\|^2 - 2\langle A^{-\frac{1}{2}}(x_k-\hat{x}), F'(x_0)(x_k-\hat{x})\rangle \\
&\quad - 2\langle A^{-\frac{1}{2}}(x_k-\hat{x}), [F'(\hat{x})-F'(x_0)](x_k-\hat{x})\rangle]
\end{aligned}
$$

$$= -2\alpha_k\left\langle A^{-\frac{1}{2}}(x_k-\hat{x}), \int_0^1 [R(\hat{x}+\theta(x_k-\hat{x}),\hat{x})-I]d\theta F'(\hat{x})(x_k-\hat{x})\right\rangle$$
$$+ \alpha_k[\alpha_k\|F(x_k)-y\|^2 - 2\langle A^{-\frac{1}{2}}(x_k-\hat{x}), F'(x_0)(x_k-\hat{x})\rangle$$
$$- 2\langle A^{-\frac{1}{2}}(x_k-\hat{x}), F'(x_0)[G(\hat{x},x_0)-I](x_k-\hat{x})\rangle].$$

The last step follows from $(\mathcal{C}_1)$. So, we have

$$\|A^{-\frac{1}{2}}(x_{k+1}-\hat{x})\|^2 - \|A^{-\frac{1}{2}}(x_k-\hat{x})\|^2$$
$$\leq 2\alpha_k\|A^{-\frac{1}{2}}(x_k-\hat{x})\|\bar{C}\int_0^1 \theta\|x_k-\hat{x}\|d\theta\|R(\hat{x},x_0)-I+I\|\|F'(x_0)(x_k-\hat{x})\|$$
$$+ \alpha_k\left[\alpha_k\left\|\int_0^1 F'(\hat{x}+\theta(x_k-\hat{x}))d\theta(x_k-\hat{x})\right\|^2 - 2\|x_k-\hat{x}\|^2 + 2\bar{C}\rho\|x_k-\hat{x}\|^2\right]$$
$$\leq 2\alpha_k\|A^{-\frac{1}{2}}(x_k-\hat{x})\|\bar{C}\int_0^1 \theta\|x_k-\hat{x}\|d\theta\|R(\hat{x},x_0)-I+I\|\|F'(x_0)(x_k-\hat{x})\|$$
$$+ \alpha_k\left[\alpha_k\left\|\int_0^1 [R(\hat{x}+\theta(x_k-\hat{x}),x_0)-I+I]d\theta F'(x_0)(x_k-\hat{x})\right\|^2 - 2\|x_k-\hat{x}\|^2 + 2\bar{C}\rho\|x_k-\hat{x}\|^2\right]$$
$$\leq \alpha_k\|A^{-\frac{1}{2}}(x_k-\hat{x})\|\frac{\bar{C}}{2}\|x_k-\hat{x}\|(1+\bar{C}\rho)\|F'(x_0)(x_k-\hat{x})\|$$
$$+ \alpha_k[\alpha_k(1+\bar{C}\rho)^2\|A^{\frac{1}{2}}(x_k-\hat{x})\|^2 - 2\|x_k-\hat{x}\|^2 + 2\bar{C}\rho\|x_k-\hat{x}\|^2].$$

Therefore by Lemma 2.1, we have

$$\|A^{-\frac{1}{2}}(x_{k+1}-\hat{x})\|^2 - \|A^{-\frac{1}{2}}(x_k-\hat{x})\|^2$$
$$\leq \alpha_k\|A^{-\frac{1}{2}}(x_k-\hat{x})\|\bar{C}\|x_k-\hat{x}\|(1+\bar{C}\rho)\|F'(x_0)(x_k-\hat{x})\| - \frac{\alpha_k}{5}\|x_k-\hat{x}\|^2$$
$$+ \alpha_k\left[\frac{(1+\bar{C}\rho)^2}{1-4\bar{C}\rho-\bar{C}^2\rho^2} + 2\bar{C}\rho - \frac{9}{5}\right]\|x_k-\hat{x}\|^2$$
$$\leq \alpha_k\|A^{-\frac{1}{2}}(x_k-\hat{x})\|\bar{C}(1+\bar{C}\rho)\|x_k-\hat{x}\|\|F'(x_0)(x_k-\hat{x})\|$$
$$- \frac{\alpha_k}{5}\|x_k-\hat{x}\|^2. \tag{2.8}$$

The last step follows from the fact that for $\bar{C}\rho \leq r_0$, we have $\frac{(1+\bar{C}\rho)^2}{1-4\bar{C}\rho-\bar{C}^2\rho^2} + 2\bar{C}\rho \leq \frac{9}{5}$. Now using the relation $2ab \leq a^2+b^2$ with $a = \sqrt{\frac{1}{5}\alpha_k}\|x_k-\hat{x}\|$ and $b = \frac{\sqrt{5\alpha_k}}{2}\bar{C}(1+\bar{C}\rho)\|A^{-\frac{1}{2}}(x_k-\hat{x})\|\|F'(x_0)(x_k-\hat{x})\|$ in (2.8), we have

$$\|A^{-\frac{1}{2}}(x_{k+1}-\hat{x})\|^2 - \|A^{-\frac{1}{2}}(x_k-\hat{x})\|^2 \leq \frac{5}{4}\bar{C}^2(1+\bar{C}\rho)^2\alpha_k\|A^{-\frac{1}{2}}(x_k-\hat{x})\|^2\|F'(x_0)(x_k-\hat{x})\|^2. \tag{2.9}$$

Now since $\bar{C}\rho \leq \min\{\sqrt{5}-2, r_0\} \leq 0.0740$, we have by (2.9)

$$\|A^{-\frac{1}{2}}(x_{k+1}-\hat{x})\|^2 - \|A^{-\frac{1}{2}}(x_k-\hat{x})\|^2 \leq 1.4418\bar{C}^2\alpha_k\|A^{\frac{1}{2}}(x_k-\hat{x})\|^2\|A^{-\frac{1}{2}}(x_k-\hat{x})\|^2.$$

Set $z_k = \|A^{-\frac{1}{2}}(x_k-\hat{x})\|$. Then

$$z_{k+1}^2 \leq (1+1.4418\bar{C}^2\alpha_k\|A^{\frac{1}{2}}(x_k-\hat{x})\|^2)z_k^2.$$

By induction

$$z_k^2 \leq \prod_{i=0}^{k-1}(1+1.4418\bar{C}^2\alpha_i\|A^{\frac{1}{2}}(x_i-\hat{x})\|^2)z_0^2. \tag{2.10}$$

The convergence of $\prod_{i=0}^{\infty}(1+1.4418\bar{C}^2\alpha_i\|A^{\frac{1}{2}}(x_i-\hat{x})\|^2)$ follows from the convergence of $\sum_{i=0}^{\infty}\alpha_i\|A^{\frac{1}{2}}(x_i-\hat{x})\|^2$. By Lemma 2.1, $\sum_{i=0}^{\infty}\alpha_i\|A^{\frac{1}{2}}(x_i-\hat{x})\|^2 < \infty$. Therefore there exist $M > 0$ such that $\sum_{i=0}^{k-1}\bar{C}^2\alpha_i\|A^{\frac{1}{2}}(x_i-\hat{x})\|^2 < M$, which implies that

$$\prod_{i=0}^{k-1}(1+1.4418\bar{C}^2\alpha_i\|A^{\frac{1}{2}}(x_i-\hat{x})\|^2) = e^{\sum_{i=0}^{k-1}\ln(1+1.4418\bar{C}^2\alpha_i\|A^{\frac{1}{2}}(x_i-\hat{x})\|^2)}$$
$$\leq e^{1.4418M}. \tag{2.11}$$

From (2.10) and (2.11), we have $z_k^2 \leq e^{1.4418M} z_0^2$. Since by $(\mathcal{C}_3)$, $z_0 = \|A^{\frac{1}{2}}(x_0 - \hat{x})\| = \|v\|$. So we have

$$z_k^2 \leq e^{1.4418M}\|v\|^2. \tag{2.12}$$

This completes the proof. □

Remark 2.3. Note that, in (2.8), one can split $-2\|x_k - \hat{x}\|^2$ into two parts, say $-c\|x_k - \hat{x}\|^2$ and $(c-2)\|x_k - \hat{x}\|^2$ such that $\frac{(1+\bar{C}\rho)^2}{1-4\bar{C}\rho-\bar{C}^2\rho^2} + 2\bar{C}\rho \leq 2 - c$. In this way one can choose a larger ρ. We choose $c = \frac{1}{5}$ for our convenience.

Theorem 2.4. *Let $(\mathcal{C})$ conditions hold and let $\bar{C}\rho < \min\{\sqrt{5} - 2, r_0\} = 0.0740$. Let x_k be as in (1.6) or (1.7). Then*

$$\|x_k - \hat{x}\| \leq \tilde{C}k^{-1/2}$$

where $\tilde{C} = \sqrt{3e^{1.4418M}}\epsilon^{-1/2}\|v\|$.

Proof. Observe that

$$\alpha_k \geq \|F'(x_0)\|^{-2}.$$

So for $\epsilon_k := \epsilon = 0.6985\|F'(x_0)\|^{-2}$, we have from Lemma 2.1, the conditions$(\mathcal{C})$ and $\bar{C}\rho \leq 0.0740$;

$$\begin{aligned}
\|x_k - \hat{x}\|^2 - \|x_{k+1} - \hat{x}\|^2 &\geq (1 - 4\bar{C}\rho - \bar{C}^2\rho^2)\alpha_k\|A^{\frac{1}{2}}(x_k - \hat{x})\|^2 \\
&\geq 0.6985\|F'(x_0)\|^{-2}\|A^{\frac{1}{2}}(x_k - \hat{x})\|^2 \\
&= \epsilon\|A^{\frac{1}{2}}(x_k - \hat{x})\|^2 \\
&= \epsilon\|F'(x_0)(x_k - \hat{x})\|^2 \\
&= \epsilon\langle F'(x_0)^*F'(x_0)(x_k - \hat{x}), x_k - \hat{x}\rangle \\
&= \epsilon\langle A(x_k - \hat{x}), x_k - \hat{x}\rangle.
\end{aligned}$$

An application of (2.1), yields

$$\begin{aligned}
\|x_k - \hat{x}\| &\leq \sqrt{3}\|A^{-\frac{1}{2}}(x_k - \hat{x})\|^{2/3}\epsilon^{-1/2}\left[\sum_{i=0}^{k-1}\|A^{-\frac{1}{2}}(x_i - \hat{x})\|^{-2/3}\right]^{-1/2} \\
&= \sqrt{3}z_k^{2/3}\epsilon^{-1/2}\left[\sum_{i=0}^{k-1} z_i^{-2/3}\right]^{-1/2}.
\end{aligned} \tag{2.13}$$

So by (2.12) and (2.13), we have

$$\begin{aligned}
\|x_k - \hat{x}\| &\leq \sqrt{3e^{1.4418M}}\epsilon^{-1/2}k^{-1/2}\|v\| \\
&\leq \tilde{C}k^{-1/2}.
\end{aligned} \tag{2.14}$$

□

3. Convergence analysis of method (1.6) and (1.7) with noisy data

In this section we consider method (1.6) and (1.7) with noisy data y^δ instead of y. As already mentioned in the introduction we assume that

$$\|y - y^\delta\| \leq \delta.$$

Precisely, we define:

$$\begin{aligned}
x_{k+1}^\delta &= x_k^\delta + \alpha_k^\delta s_k^\delta \quad (k = 0, 1, 2, \ldots) \\
s_k^\delta &= -F'(x_0)^*(F(x_k^\delta) - y^\delta) \\
\alpha_k^\delta &= \frac{\|s_k^\delta\|^2}{\|F'(x_0)s_k^\delta\|^2}
\end{aligned} \tag{3.1}$$

and

$$\begin{aligned}
x_{k+1}^\delta &= x_k^\delta + \alpha_k^\delta s_k^\delta \quad (k = 0, 1, 2, \ldots) \\
s_k^\delta &= -F'(x_0)^*(F(x_k^\delta) - y^\delta) \\
\alpha_k^\delta &= \frac{\|F(x_k^\delta) - y^\delta\|^2}{\|s_k^\delta\|^2},
\end{aligned} \tag{3.2}$$

instead of x_k in (1.6) and (1.7), respectively. We will use the following assumption together with the assumptions $(\mathcal{C})$:

$(\mathcal{C}_4)$ F satisfies the local property

$$\|F(u)-F(v)-F'(x_0)(u-v)\| \le \eta\|F(u)-F(v)\| \quad (\eta \le \frac{1}{2}) \tag{3.3}$$

for all $u, v \in B(x_0, \rho)$.

It was shown in [4–8] (for linear ill-posed problems) that the steepest descent method converges in the case of exact data, but due to the instability of the steepest method it is impossible to use a-priori parameter choice strategies as stopping criteria. Therefore, a-priori strategy is used in the literature [13] for stopping (3.1) and (3.2). As already mentioned in the introduction, a-priori stopping rules are used for steepest descent method in the literature to prove the convergence of (3.1) and (3.2) to $\hat{x}$ but no error estimate for $\|x_k^\delta - \hat{x}\|$ was given (as far as the authors are known). In this section we propose a discrepancy principle for method (3.1) and (3.2).

3.1. Discrepancy principle

Proposition 3.1. *Let $(\mathcal{C})$ conditions hold. Let x_k^δ be as in (3.1) or (3.2). Then, a sufficient condition for x_{k+1}^δ, to be a better approximation of $\hat{x}$ is that,*

$$\|F(x_k^\delta)-y^\delta\| > \tau\delta \tag{3.4}$$

where

$$\tau > 2\frac{(1+\eta)}{1-2\eta} > 2. \tag{3.5}$$

In particular, if $x_k^\delta \in D(F)$ and (3.4) holds for all $0 \le k < k_*$ with τ as in (3.5), then

$$k_*(\tau\delta)^2 \le \sum_{k=0}^{k_*-1}\|F(x_k^\delta)-y^\delta\|^2 \le \frac{\tau\|F'(x_0)\|^2}{(1-2\eta)\tau-2(1+\eta)}\|x_0-\hat{x}\|^2. \tag{3.6}$$

Proof. Using (3.1) or (3.2), we have

$$\begin{aligned}
\|x_{k+1}^\delta-\hat{x}\|^2-\|x_k^\delta-\hat{x}\|^2 &= -2\alpha_k^\delta\langle x_k^\delta-\hat{x}, F'(x_0)^*(F(x_k^\delta)-y^\delta)\rangle + {\alpha_k^\delta}^2\|F'(x_0)^*(F(x_k^\delta)-y^\delta)\|^2\\
&= 2\alpha_k^\delta\langle F(x_k^\delta)-y^\delta-F'(x_0)(x_k^\delta-\hat{x}), F(x_k^\delta)-y^\delta\rangle\\
&\quad + \alpha_k^\delta[\alpha_k^\delta\|F'(x_0)^*(F(x_k^\delta)-y^\delta)\|^2 - 2\|F(x_k^\delta)-y^\delta\|^2]\\
&\le 2\alpha_k^\delta\|F(x_k^\delta)-F(\hat{x})+y-y^\delta-F'(x_0)(x_k^\delta-\hat{x})\|\|F(x_k^\delta)-y^\delta\|\\
&\quad + \alpha_k^\delta[\alpha_k^\delta\|F'(x_0)^*(F(x_k^\delta)-y^\delta)\|^2 - 2\|F(x_k^\delta)-y^\delta\|^2].
\end{aligned} \tag{3.7}$$

So by $(\mathcal{C}_4)$, we have by (3.7)

$$\begin{aligned}
\|x_{k+1}^\delta-\hat{x}\|^2-\|x_k^\delta-\hat{x}\|^2 &\le 2\alpha_k^\delta(\eta\|F(x_k^\delta)-F(\hat{x})\|+\delta)\|F(x_k^\delta)-y^\delta\|\\
&\quad + \alpha_k^\delta[\alpha_k^\delta\|F'(x_0)^*(F(x_k^\delta)-y^\delta)\|^2 - 2\|F(x_k^\delta)-y^\delta\|^2]\\
&\le 2\alpha_k^\delta[\eta\|F(x_k^\delta)-y^\delta\|+(1+\eta)\delta]\|F(x_k^\delta)-y^\delta\|\\
&\quad + \alpha_k^\delta[\alpha_k^\delta\|F'(x_0)^*(F(x_k^\delta)-y^\delta)\|^2 - 2\|F(x_k^\delta)-y^\delta\|^2]\\
&= \alpha_k^\delta(2\eta-1)\|F(x_k^\delta)-y^\delta\|^2\\
&\quad + \alpha_k^\delta 2(1+\eta)\delta\|F(x_k^\delta)-y^\delta\|\\
&\quad + \alpha_k^\delta[\alpha_k^\delta\|F'(x_0)^*(F(x_k^\delta)-y^\delta)\|^2 - \|F(x_k^\delta)-y^\delta\|^2].
\end{aligned}$$

In both methods, i.e., (3.1) and (3.2), $\alpha_k^\delta\|F'(x_0)^*(F(x_k^\delta)-y^\delta)\|^2 \le \|F(x_k^\delta)-y^\delta\|^2$. Therefore, we have

$$\|x_{k+1}^\delta-\hat{x}\|^2-\|x_k^\delta-\hat{x}\|^2 \le \alpha_k^\delta[(2\eta-1)\|F(x_k^\delta)-y^\delta\|^2+2(1+\eta)\delta\|F(x_k^\delta)-y^\delta\|], \tag{3.8}$$

so, by (3.4), we have

$$\|x_{k+1}^\delta-\hat{x}\|^2-\|x_k^\delta-\hat{x}\|^2 \le \alpha_k^\delta\left((2\eta-1)+2\frac{(1+\eta)}{\tau}\right)\|F(x_k^\delta)-y^\delta\|^2 < 0. \tag{3.9}$$

Now since $\alpha_k^\delta \ge \|F'(x_0)\|^{-2}$, we have by (3.9)

$$\|F'(x_0)\|^{-2}\left((1-2\eta)-2\frac{(1+\eta)}{\tau}\right)\|F(x_k^\delta)-y^\delta\|^2 \le \|x_k^\delta-\hat{x}\|^2-\|x_{k+1}^\delta-\hat{x}\|^2. \tag{3.10}$$

Adding the inequality (3.10) for k from 0 through $k_* - 1$, we obtain

$$\|F'(x_0)\|^{-2}\left((1-2\eta) - 2\frac{(1+\eta)}{\tau}\right)\sum_{k=0}^{k_*-1}\|F(x_k^\delta) - y^\delta\|^2 \leq \|x_0 - \hat{x}\|^2 - \|x_{k_*}^\delta - \hat{x}\|^2. \tag{3.11}$$

This completes the proof. □

Remark 3.2. Thus (3.11) implies that, for $y^\delta \neq y$, there must be a unique index k_* such that (3.4) holds for all $k < k_*$ but is violated at $k = k_*$ (see also [9, p. 282]).

Next Lemma is used to prove our main result in this section.

Lemma 3.3. *Let $\bar{C}\rho < \frac{2(\tau-2)}{\tau}$. Then $\delta \leq (1 - \frac{\bar{C}}{2}\|x_k^\delta - \hat{x}\|)\|F'(\hat{x})(x_k^\delta - \hat{x})\|$ for all $0 < k \leq k_*$.*

Proof. By the definition of k_*, we have for $k \leq k_*$;

$$\begin{aligned}
\tau\delta &< \|F(x_k^\delta) - y^\delta\| \\
&\leq \|F(x_k^\delta) - F(\hat{x})\| + \|y - y^\delta\| \\
&\leq \left\|\int_0^1 F'(\hat{x} + \theta(x_k^\delta - \hat{x}))d\theta(x_k^\delta - \hat{x})\right\| + \delta \\
&\leq \left\|\int_0^1 [R(\hat{x} + \theta(x_k^\delta - \hat{x}), \hat{x}) - I + I]d\theta F'(\hat{x})(x_k^\delta - \hat{x})\right\| + \delta \\
&\leq \int_0^1 [\bar{C}\|\theta(x_k^\delta - \hat{x})\| + 1]d\theta\|F'(\hat{x})(x_k^\delta - \hat{x})\| + \delta \\
&\leq \left(1 + \frac{\bar{C}}{2}\|x_k^\delta - \hat{x}\|\right)\|F'(\hat{x})(x_k^\delta - \hat{x})\| + \delta \\
&\leq \left(1 + \frac{\bar{C}}{2}\rho\right)\|F'(\hat{x})(x_k^\delta - \hat{x})\| + \delta.
\end{aligned}$$

The last step follows from (3.9) (i.e., $\|x_k^\delta - \hat{x}\| < \|x_0 - \hat{x}\| < \rho$). Thus, we have

$$\begin{aligned}
\delta &\leq \frac{1 + \frac{\bar{C}\rho}{2}}{\tau - 1}\|F'(\hat{x})(x_k^\delta - \hat{x})\| \\
&\leq \left(1 - \frac{\bar{C}\rho}{2}\right)\|F'(\hat{x})(x_k^\delta - \hat{x})\| \\
&\leq \left(1 - \frac{\bar{C}}{2}\|x_k^\delta - \hat{x}\|\right)\|F'(\hat{x})(x_k^\delta - \hat{x})\|.
\end{aligned}$$

This completes the proof. □

Let $\Omega := \|F'(x_0)\|^{-2}\left((1-2\eta) - 2\frac{(1+\eta)}{\tau}\right)$.

Theorem 3.4. *Let $(\mathcal{C})$ conditions hold and let $\bar{C}\rho < \min\{\frac{2(\tau-2)}{\tau}, \frac{2}{m\sqrt{\Omega}}, 1\}$. Let x_{k+1}^δ be as in (3.1) or (3.2). Then for $0 \leq k < k_*$,*

$$\|x_{k+1}^\delta - \hat{x}\| = \begin{cases} O(q^{\frac{k+1}{2}}) & if \quad \delta < q^{k+1} \\ O(\delta^{\frac{1}{2}}) & if \quad q^{k+1} \leq \delta \end{cases} \tag{3.12}$$

where $q := \max\{1 - \frac{\bar{C}^2\Omega}{4}\|F'(\hat{x})(x_i^\delta - \hat{x})\|^2 : i = 0, 1, 2, \ldots k\}$.

Proof. Since F is Fréchet differentiable at $\hat{x}$,

$$\|F(x_k^\delta) - F(\hat{x}) - F'(\hat{x})(x_k^\delta - \hat{x})\| = \left\|\int_0^1 (F'(\hat{x} + \theta(x_k^\delta - \hat{x})) - F'(\hat{x}))d\theta(x_k^\delta - \hat{x})\right\|. \tag{3.13}$$

So by $(\mathcal{C}_2)$,

$$\begin{aligned}
\|F(x_k^\delta) - F(\hat{x}) - F'(\hat{x})(x_k^\delta - \hat{x})\| &= \left\|\int_0^1 (R(\hat{x} + \theta(x_k^\delta - \hat{x}), \hat{x}) - I)F'(\hat{x})(x_k^\delta - \hat{x})d\theta\right\| \\
&\leq \frac{\bar{C}}{2}\|x_k^\delta - \hat{x}\|\|F'(\hat{x})(x_k^\delta - \hat{x})\|.
\end{aligned}\tag{3.14}$$

Using (3.14), we have,

$$\begin{aligned}\|F'(\hat{x})(x_k^\delta-\hat{x})\| - \frac{\bar{C}}{2}\|x_k^\delta-\hat{x}\|\|F'(\hat{x})(x_k^\delta-\hat{x})\| &\\ \leq \|F'(\hat{x})(x_k^\delta-\hat{x})\| - \|F(x_k^\delta)-F(\hat{x})-F'(\hat{x})(x_k^\delta-\hat{x})\| &\\ \leq \|F(x_k^\delta)-F(\hat{x})\|. &\end{aligned} \tag{3.15}$$

Thus, by (3.15) and Lemma 3.3 we have

$$\begin{aligned}\|F(x_k^\delta)-y^\delta\| &\geq \|F(x_k^\delta)-F(\hat{x})\|-\delta\\ &\geq \|F'(\hat{x})(x_k^\delta-\hat{x})\| - \frac{\bar{C}}{2}\|x_k^\delta-\hat{x}\|\|F'(\hat{x})(x_k^\delta-\hat{x})\|-\delta>0.\end{aligned} \tag{3.16}$$

Thus,

$$\begin{aligned}\|F(x_k^\delta)-y^\delta\|^2 \geq\ & \|F'(\hat{x})(x_k^\delta-\hat{x})\|^2\\ &+\left(\frac{\bar{C}}{2}\right)^2\|x_k^\delta-\hat{x}\|^2\|F'(\hat{x})(x_k^\delta-\hat{x})\|^2\\ &+\delta^2+\bar{C}\|x_k^\delta-\hat{x}\|\|F'(\hat{x})(x_k^\delta-\hat{x})\|\delta\\ &-\bar{C}\|x_k^\delta-\hat{x}\|\|F'(\hat{x})(x_k^\delta-\hat{x})\|^2\\ &-2\delta\|F'(\hat{x})(x_k^\delta-\hat{x})\|.\end{aligned} \tag{3.17}$$

So by (3.17) and (3.10), we have

$$\begin{aligned}\|x_{k+1}^\delta-\hat{x}\|^2 \leq\ & \left(1-\frac{\bar{C}^2\Omega}{4}\|F'(\hat{x})(x_k^\delta-\hat{x})\|^2\right)\|x_k^\delta-\hat{x}\|^2\\ &+(\bar{C}\rho-1)\Omega\|F'(\hat{x})(x_k^\delta-\hat{x})\|^2\\ &-\Omega\delta^2-\Omega\bar{C}\|x_k^\delta-\hat{x}\|\|F'(\hat{x})(x_k^\delta-\hat{x})\|\delta\\ &+2\Omega\delta\|F'(\hat{x})(x_k^\delta-\hat{x})\|\\ \leq\ & \left(1-\frac{\bar{C}^2\Omega}{4}\|F'(\hat{x})(x_k^\delta-\hat{x})\|^2\right)\|x_k^\delta-\hat{x}\|^2\\ &+2\Omega m\rho\delta.\end{aligned}$$

Therefore we have,

$$\|x_{k+1}^\delta-\hat{x}\|^2 \leq q_k\|x_k^\delta-\hat{x}\|^2+L\delta$$

where $q_k=1-\frac{\bar{C}^2\Omega}{4}\|F'(\hat{x})(x_k^\delta-\hat{x})\|^2$ and $L=2\Omega m\rho$.

Note that $q_k=1-\frac{\bar{C}^2\Omega}{4}\|F'(\hat{x})(x_k^\delta-\hat{x})\|^2<1$. Then,

$$\begin{aligned}\|x_{k+1}^\delta-\hat{x}\|^2 &\leq q^{k+1}\|x_0^\delta-\hat{x}\|^2+q^kL\delta+\ldots+qL\delta+L\delta\\ &\leq q^{k+1}\rho^2+\frac{L\delta}{1-q}.\end{aligned} \tag{3.18}$$

This completes the proof. □

3.2. Convergence rate result for steepest descent method and minimal error method with noisy data

In this section we consider the steepest descent method and minimal error method with noisy data and obtained a convergence rate result which is not available in the literature. The steepest descent method and minimal error method with noisy data are defined by

$$\begin{aligned}x_{k+1}^\delta &= x_k^\delta+\alpha_k^\delta s_k^\delta\quad (k=0,1,2,\ldots)\\ s_k^\delta &= -F'(x_k^\delta)^*(F(x_k^\delta)-y^\delta)\\ \alpha_k^\delta &= \frac{\|s_k^\delta\|^2}{\|F'(x_k^\delta)s_k^\delta\|^2}\end{aligned} \tag{3.19}$$

and

$$\begin{aligned} x_{k+1}^\delta &= x_k^\delta + \alpha_k^\delta s_k^\delta \quad (k = 0, 1, 2, \ldots) \\ s_k^\delta &= -F'(x_k^\delta)^*(F(x_k^\delta) - y^\delta) \\ \alpha_k^\delta &= \frac{\|F(x_k^\delta) - y^\delta\|^2}{\|s_k^\delta\|^2}, \end{aligned} \tag{3.20}$$

respectively. We have the following convergence rate result.

Theorem 3.5. *Let $(\mathcal{C})$ conditions hold and let $\bar{C}\rho < \min\{\frac{2(\tau-2)}{\tau}, \frac{2}{m\sqrt{\omega}}, 1\}$. Let x_{k+1}^δ be as in (3.19) or (3.20). Then for $0 \le k < k_*$,*

$$\|x_{k+1}^\delta - \hat{x}\| = \begin{cases} O(q^{\frac{k+1}{2}}) & if \quad \delta < q^{k+1} \\ O(\delta^{\frac{1}{2}}) & if \quad q^{k+1} \le \delta \end{cases} \tag{3.21}$$

where $q := \max\{1 - \frac{\bar{C}^2\omega}{4}\|F'(\hat{x})(x_i^\delta - \hat{x})\|^2 : i = 0, 1, 2, \ldots k\}$ with $\omega := \|F'(x_k^\delta)\|^{-2}\Big((1-2\eta) - 2\frac{(1+\eta)}{\tau}\Big)$.

Proof. Simply follow the proof of Theorem 3.4. □

4. Example

In this section we present an academic example which satisfies the assumptions $(\mathcal{C}_1)$ and $(\mathcal{C}_2)$.

Example 4.1. Consider a nonlinear operator equation F: $L^2[0, 1] \to L^2[0, 1]$ defined by

$$F(u) := (\arctan(u))^2. \tag{4.22}$$

The Fréchet derivative of F is

$$F'(u)w = \frac{2arctan(u)}{1+u^2}w.$$

If $u(x)$ vanishes on a set of positive Lebesgue measure, then $F'(u)$ is not boundedly invertible. If $u \in C[0, 1]$ vanishes even at one point x_0, then $F'(u)$ is not boundedly invertible in $L^2[0, 1]$.

Note that

$$F'(\hat{u})w = F'(u_0)G(\hat{u}, u_0)w,$$

and

$$F'(u)w = R(u, u_0)F'(u_0)w$$

with $G(\hat{u}, u_0) = \frac{1+u_0^2}{1+\hat{u}^2}\frac{arctan(\hat{u})}{arctan(u_0)}$ and $R(u, u_0) = \frac{1+u_0^2}{1+u^2}\frac{arctan(u)}{arctan(u_0)}$, respectively. Further, for $u_0 \ne 0$,

$$\|G(\hat{u}, u_0) - I\| \le \left[\frac{1}{\|arctan(u_0)\|} + 2\max\{\|\hat{u}\|, \|u_0\|\}\right]\|\hat{u} - u_0\|$$

and

$$\|R(u, u_0) - I\| \le \left[\frac{1}{\|arctan(u_0)\|} + 2\max\{\|u\|, \|u_0\|\}\right]\|u - u_0\|.$$

That is, assumptions $(\mathcal{C}_1)$ and $(\mathcal{C}_2)$ are satisfied.

Acknowledgment

One of the authors, Sabari, thanks National Institute of Technology Karnataka, India, for the financial support, grant No. Ph.D/ Fellowship/ 2014-15/ E1.

References

[1] H. Brakhage, On ill-posed problems and the method of conjugate gradients, Inverse and Ill-Posed Problems, Academic Press, New York, 1987, pp. 165–175.
[2] B. Eicke, A.K. Louis, R. Plato, The instability of some gradient methods for ill-posed problems, Numer. Math. 58 (1990) 129–134.
[3] G.H. Goulb, D.P. O'Leary, Some history of the conjugate gradient method and Lanczos algorithms: 1948–1976, SIAM Rev. 31 (1989) 50–102.
[4] W.J. Kammerer, M.Z. Nashed, Steepest descent for singular linear operators with non closed range, Applicable Anal. 1 (1971) 143–159.
[5] W.J. Kammerer, M.Z. Nashed, On the convergence of the conjugate gradient method for singular linear operators, SIAM J. Numer. Anal. 9 (1972) 165–181.
[6] J.T. King, A minimal error conjugate gradient method for ill-posed problems, J. Optim. Theory Appl. 60 (1989) 297–304.
[7] L.J. Lardy, A class of iterative methods of conjugate gradient type, Numer. Funct. Anal. Optim. 11 (1990) 283–302.

[8] A.K. Louis, Convergence of the conjugate gradient method for compact operators, Inverse and Ill-Posed Problems, Academic Press, New York, 1987, pp. 177–183.
[9] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems; Mathematics and Its Applications, Kluwer Academic Publishers, Dordrecht, 2000.
[10] S.F. Gilyazov, Iterative solution methods for inconsistent linear equations with non self-adjoint operators, Moscow Univ. Comp. Math. Cyb. 1 (1997) 8–13.
[11] A. Neubauer, O. Scherzer, A convergence rate result for a steepest descent method and a minimal error method for the solution of nonlinear ill-posed problems, Z. Anal. Anwend. 14 (2) (1995) 369–377.
[12] R. Ramlau, TIGRA—an iterative algorithm for regularizing nonlinear ill-posed problems, Inverse Problems 19 (2003) 433–465.
[13] O. Scherzer, A convergence analysis of a method of steepest descent and two-step algorithm for nonlinear ill-posed problems, Numer. Funct. Anal. Optim. 17 (1–2) (1996) 197–214.
[14] V. Vasin, Irregular nonlinear operator equations: Tikhonov's regularization and iterative approximation, J. Inverse Ill Posed Problem 21 (2013) 109–123.