A convergence analysis for Tikhonov regularization of nonlinear ill-posed problems

This article has been downloaded from IOPscience. Please scroll down to see the full text article.
1999 Inverse Problems 15 1087
(http://iopscience.iop.org/0266-5611/15/4/316)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 134.109.41.2
The article was downloaded on 16/01/2012 at 15:09

Please note that terms and conditions apply.
A convergence analysis for Tikhonov regularization of nonlinear ill-posed problems

Jin Qi-nian
Institute of Mathematics, Nanjing University, Nanjing 210008, People’s Republic of China
E-mail: galgebra@nju.edu.cn

Received 28 April 1998, in final form 19 January 1999

Abstract. In this paper we consider the a posteriori parameter choice strategy proposed by Scherzer et al in 1993 for the Tikhonov regularization of nonlinear ill-posed problems and obtain some results on the convergence and convergence rate of Tikhonov regularized solutions under suitable assumptions. Finally we present some illustrative examples.

1. Introduction

This paper concerns the approximate resolution of the nonlinear equation

\[ F(x) = y_0 \]  

by means of Tikhonov regularization, where \( F \) is a nonlinear operator with domain \( D(F) \) in the Hilbert space \( X \) and with its range \( R(F) \) in the Hilbert space \( Y \), and the data \( y_0 \) are attainable, i.e. \( y_0 \in R(F) \). Throughout this paper it is assumed that \( F \) is weakly closed, continuous and Fréchet differentiable; the Fréchet derivative of \( F \) at \( x \in D(F) \) and its adjoint are denoted by \( F'(x) \) and \( F'(x)^* \), respectively. The interest of this paper is confined to the case that problem (1.1) is ill-posed, i.e. the solution of (1.1) lacks continuous dependence on the right-hand side; the readers can consult [3, 5] for a number of important inverse problems in natural science leading to such a case.

By assuming \( y_\delta \) to be the only available approximation of \( y_0 \) satisfying

\[ \| y_\delta - y_0 \| \leq \delta \]  

with a given noise level \( \delta > 0 \), now the reconstruction of the solution of (1.1) comes into being. Tikhonov regularization can be applied to pursue this task and the solution \( x_\alpha \) of the minimization problem

\[ \min_{x \in D(F)} \{ \| F(x) - y_\delta \|^2 + \alpha \| x - x^* \|^2 \} \]  

(1.3)

can be used as an approximate solution of (1.1), where \( \alpha > 0 \) is the regularization parameter and \( x^* \in X \) is an a priori guess of the exact solution. Under a suitable a priori choice of \( \alpha \), \( x_\alpha \) can be guaranteed to converge to an \( x^* \)-minimum-norm solution (\( x^* \)-MNS) \( x_0 \) of (1.1), i.e. converge to an element \( x_0 \in X \) with the property

\[ F(x_0) = y_0 \quad \text{and} \quad \| x_0 - x^* \| = \min_{x \in D(F)} \{ \| x - x^* \| : F(x) = y_0 \} \]

0266-5611/99/041087+12$30.00 © 1999 IOP Publishing Ltd
and furthermore, the estimate of \( \|x_0^\delta - x_0\| \) can be derived if \( x_0 \) has some kind of source-wise representation, see [2, 10].

Although it gives some interesting insights into the Tikhonov regularized solution \( x_0^\delta \), the \textit{a priori} choice strategy is useless in practice since it depends on the smoothness on \( x_0 - x^* \), which is difficult to check in general. Thus a wrong guess of the smoothness will lead to a bad choice of the regularization parameter and consequently to a bad approximation of \( x_0 \).

Among these rules, the most attractive is the one proposed by Scherzer et al in 1993 (see [12]), and the regularization parameter \( \alpha \) is chosen as the root of the nonlinear equation

\[
\alpha(F(x_0^\delta) - y_\delta, (\alpha I + F'(x_0^\delta)F'(x_0^\delta)^{-1}(F(x_0^\delta) - y_\delta)) = c\delta^2. \tag{1.4}
\]

Some analyses have been given on this rule in [12]. Unfortunately the results therein are not really applicable in practice since very restrictive conditions, most of which cannot be checked for concrete problems at all, have been exerted on \( F \). Some attempts have been made in [9] to retrieve this strategy from such embarrassment, and the validity of all the results in [12] has been proved under suitable assumptions, and hence a theoretical justification of this rule has been established. In this paper we continue such research and try to obtain some useful results under some conditions which can be viewed as a replenishment of those in [9].

The organization of this paper is as follows. In section 2 we recall the existing results and give some comments on the limitation of the conditions needed in the literature. Then we prove a convergence result in section 3 and derive a result on convergence rate in section 4 under certain conditions. Finally we present some examples in section 5 to illustrate the conditions required in the foregoing sections.

2. Some existing results

Before continuing our effort to study rule (1.4), let us recall the existing results on this strategy and give some comments on the conditions. We state rule (1.4) in the following general form.

**Rule 2.1.** Let \( c \geq 1 \) be a given constant and \( x^* \in D(F) \).

(i) If \( \|F(x^*) - y_\delta\|^2 \leq c\delta^2 \), then choose \( \alpha = \infty \), i.e. take \( x^* \) as approximation;

(ii) If \( \|F(x^*) - y_\delta\|^2 > c\delta^2 \), then choose \( \alpha := \alpha(\delta) \) as the root of equation (1.4).

The justification of rule 2.1 can be confirmed under certain conditions. If we assume that (1.1) has an \( x^*\)-MNS \( x_0 \) such that

\[
B_p(x_0) \subset D(F) \tag{2.1}
\]

with some number \( p > 3\|x_0 - x^*\| \), and there exists a constant \( K_0 \) such that for all \( x, z \in B_p(x_0) \) and \( v \in X \), there is \( k(x, z, v) \in X \) such that

\[
(F'(x) - F'(z))v = F'(z)k(x, z, v) \tag{2.2}
\]

with

\[
\|k(x, z, v)\| \leq K_0\|x - z\\|\|v\| \tag{2.3}
\]

then rule 2.1 is well defined provided \( c \geq 2 \) and \( 2K_0\|x_0 - x^*\| < 1 \) (please refer to lemma 2.1 in [9] or theorem 3.9 in [12] under more complicated requirements).

With the \( \alpha := \alpha(\delta) \) chosen by rule 2.1 we hope to obtain the approximation property of \( x_0^\delta \). This question was first considered in [12] under so strong conditions that we are not
Tikhonov regularization of nonlinear ill-posed problems

Sure whether the results therein can be applied when we handle concrete problems, although the numerical results in [12] give a convincing illustration. Now it is natural for us to ask whether the results in [12] are still valid under assumptions which can be really checked. A first positive answer was given in [9] by theorem 1.3 and theorem 1.4 merely under conditions (2.1), (2.2) and (2.3), and is condensed in the following result.

**Theorem 2.1.** Let (2.1)–(2.3) hold, \( c > 9 \) and let \( \alpha(\delta) \) be determined by rule 2.1.

(i) If \( x_0 \) is the unique \( x^*\)-MNS of (1.1) and \( 2K_0 \| x_0 - x^* \| < 1 \), then

\[
\lim_{\delta \to 0} \| x_{\alpha(\delta)} - x_0 \| = 0.
\]

(ii) If \( 6K_0 \| x_0 - x^* \| \leq 1 \) and if there is a \( 0 < \nu \leq 2 \) and an element \( \omega \in \mathcal{N}(F'(x_0)) \perp \subset X \) such that \( x_0 - x^* = (F'(x_0)^* F'(x_0))^{\nu/2} \omega \), then there is a constant \( C_\nu \) depending on \( \nu \) only such that

\[
\| x_{\alpha(\delta)} - x_0 \| \leq C_\nu \| \omega \|^{1/(1+\nu)} \delta^{\nu/(1+\nu)}.
\]

This result shows that Tikhonov regularization combining with rule 2.1 defines a regularization method of optimal order for each \( 0 < \nu \leq 2 \), and it explains the reason why rule 2.1 has elegant performance for the numerical examples in [12] which does not satisfy the assumptions therein. In [8] we give a further study on rule 2.1 under (2.2) and (2.3) and obtain the optimality in the sense of [12, definition 1.1] by restricting the spectra of \( F'(x_0)^* F'(x_0) \).

We assume there exists a decreasing sequence \( \{ \lambda_k \} \subset \sigma(F'(x_0)^* F'(x_0)) = \sigma(F'(x_0) F'(x_0)^*) \) satisfying

\[
\lim_{k \to \infty} \lambda_k = 0 \quad \text{and} \quad \frac{\lambda_k}{\lambda_{k+1}} \leq C \quad \text{for all} \quad k \quad (2.4)
\]

with a constant \( C > 1 \), then we obtain (see [8, theorem 4]).

**Theorem 2.2.** Let (2.1)–(2.3) and (2.4) hold, \( 6K_0 \| x_0 - x^* \| \leq 1 \) and let \( \alpha(\delta) \) be determined by rule 2.1. Then there exists a constant \( C_\nu \) and a positive number \( \delta_0 \) such that for all \( 0 < \delta \leq \delta_0 \) there holds

\[
\sup \{ \| x_{\alpha(\delta)} - x_0 \| : \| y_0 - y_0 \| \leq \delta \} \leq C_\nu \hat{\psi}_{y_0}(\delta)
\]

where \( \hat{\psi}_{y_0}(\delta) \) is the optimal convergence rate for \( y_0 \) defined by

\[
\hat{\psi}_{y_0}(\delta) := \sup \{ \inf \{ \| x_{\alpha(\delta)} - x_0 \| : \alpha > 0 \} : \| y_0 - y_0 \| \leq \delta \}.
\]

Let us give some comments on (2.2) and (2.3). Although they are not too restrictive and can be verified for many concrete problems, there are some critical cases in which (2.2) and (2.3) are violated. We give an illustration by the following example.

**Example 2.1.** Consider the problem of estimating the diffusion coefficient \( a \) in the two-point boundary value problem

\[
\begin{align*}
- (au_t)_t &= f & \text{in } (0, 1) \\
u(0) &= g_0 & u(1) = g_1
\end{align*}
\]

from noisy data \( u_\delta \) of the state \( u_0 := u(a_0) \), where \( f \in L^2, g_0, g_1 \) are real numbers and \( a_0 \) is the sought solution. We can define the nonlinear operator \( F \) by

\[
F : D(F) := \{ a \in H^1 : a(t) \geq \mu > 0 \ \text{a.e.} \} \subset H^1[0, 1] \mapsto L^2[0, 1]
\]

\[
a \mapsto F(a) := u(a)
\]
where \(u(a)\) is the unique solution of (2.5). It is well known that (see [12]) if \(a_0\) admits the property
\[
|u_r(a_0)(t)| \geq \kappa \quad \text{for all } t \in [0, 1]
\]
with some positive constant \(\kappa\), then the Fréchet derivative of \(F\) satisfies (2.2) and (2.3) in a neighbourhood of \(a_0\). However, if (2.6) is violated, which covers the problems with homogeneous boundary conditions, \(F\) does not satisfy (2.2) and (2.3) again. Under such circumstances, we now wonder whether the assertions in theorem 2.1 are still valid.

The above example shows that it is necessary to derive some useful conclusions under some conditions different from (2.2) and (2.3). We will do this in the next two sections.

3. Convergence criterion

We begin this section by first discussing the justification of rule 2.1 without (2.2) and (2.3). Obviously, we only need to consider the case \(\alpha(\delta)\) for the \(\alpha\).

Let (3.1) and (3.2) hold, then
\[
\text{Theorem 3.1. For details, please refer to [7] for an analogous discussion by borrowing the idea in the proof of [2, theorem 2.1].}
\]

We assume
\[
\begin{align*}
x \mapsto F(x) \text{ and } x \mapsto F'(x) \text{ are continuous on } D(F).
\end{align*}
\]

Suppose \(c > 1\) and set
\[
\begin{align*}
\alpha_0 & := \frac{(c - 1)\delta^2}{\|x_0 - x^*\|^2} \\
\rho(\alpha) & := \alpha(F(x_0^\delta) - y_\delta, (\alpha I + F'(x_0^\delta)F'(x_0^\delta)^*)^{-1}(F(x_0^\delta) - y_\delta))
\end{align*}
\]

then the definition of \(x_0^\delta\) gives \(\rho(\alpha_0) \leq \|F(x_0^\delta) - y_\delta\|^2 \leq \delta^2 + \alpha_0\|x_0 - x^*\|^2 = c\delta^2\). From the definition of \(x_0^\delta\) it follows \(x_0^\delta \rightarrow x^*\) as \(\alpha \rightarrow \infty\). By letting \(B_0^\delta := F'(x_0^\delta)F'(x_0^\delta)^*\), then (3.1) implies \(\|B_0^\delta\|\) is bounded and \(\|(\alpha I + B_0^\delta)^{-1}B_0^\delta\| \rightarrow 0\) as \(\alpha \rightarrow \infty\). Therefore from
\[
|\rho(\alpha) - \|F(x_0^\delta) - y_\delta\|^2| = (F(x_0^\delta) - y_\delta, (\alpha I + B_0^\delta)^{-1}B_0^\delta(F(x_0^\delta) - y_\delta)) \leq \|B_0^\delta\|\|F(x_0^\delta) - y_\delta\|^2
\]

it gives \(\lim_{\alpha \rightarrow \infty} \rho(\alpha) = \lim_{\alpha \rightarrow \infty} \|F(x_0^\delta) - y_\delta\|^2 = \|F(x^*) - y_\delta\|^2 > c\delta^2\). Hence we can conclude the existence of an \(\alpha(\delta) \geq \alpha_0\) satisfying (1.4) if \(\rho(\alpha)\) is continuous with respect to \(\alpha\) on \([\alpha_0, \infty)\). Thanks to (3.1), we need only to show the continuity of the mapping \(\alpha \mapsto x_0^\delta\) for \(\alpha \in [\alpha_0, \infty)\). This can be guaranteed if

the minimization problem (1.3) has a unique solution \(x_0^\delta\) for each \(\alpha \geq \alpha_0\).

For details, please refer to [7] for an analogous discussion by borrowing the idea in the proof of [2, theorem 2.1].

Now we can give the convergence of \(x_{\alpha(\delta)}^\delta\).

**Theorem 3.1.** Let (3.1) and (3.2) hold, \(c > 1\) and let \(F\) be weakly closed. If (1.1) has a unique \(x^*-\mbox{MNS} x_0\), then
\[
\lim_{\delta \rightarrow 0} x_{\alpha(\delta)}^\delta = x_0
\]
for the \(\alpha(\delta)\) determined by rule 2.1.

**Proof.** The proof can be carried out by considering the following three different cases.

Suppose that there is a sequence \(\delta_k\) such that \(\delta_k \rightarrow 0\) and \(\alpha(\delta_k) \rightarrow \infty\) as \(k \rightarrow \infty\). Since \(x_{\alpha(\delta)}^\delta \rightarrow x^*\) as \(k \rightarrow \infty\), it follows from the definition of \(\alpha(\delta_\ell)\) and (3.1) that
\[
\begin{align*}
0 &= \lim_{k \rightarrow \infty} \alpha(\delta_k)(F(x_{\alpha(\delta_k)}^\delta) - y_{\delta_k}, (\alpha(\delta_k) I + F'(x_{\alpha(\delta_k)}^\delta)F'(x_{\alpha(\delta_k)}^\delta)^*)^{-1}(F(x_{\alpha(\delta_k)}^\delta) - y_{\delta_k})) \\
&= \lim_{k \rightarrow \infty} \|F(x_{\alpha(\delta_k)}^\delta) - y_{\delta_k}\|^2 = \|F(x^*) - y_0\|^2
\end{align*}
\]
which implies \( x^* \) is a solution of (1.1), hence \( x_0 = x^* \) and \( x_{\alpha(k)} \to x_0 \).

Assume next that there is a sequence \( \delta_k \) such that \( \delta_k \to 0 \) and \( \alpha(\delta_k) \to \beta \) with a positive number \( \beta < \infty \) as \( k \to \infty \). By letting \( x_\beta \) be a solution of the minimization problem (1.3) with \( y_\beta \) and \( \alpha \) replaced by \( y_0 \) and \( \beta \) respectively, then the definition of \( x_{\alpha(k)} \) gives
\[
\| F(x_{\alpha(k)}) - y_{\delta_k} \|^2 + \alpha(\delta_k) \| x_{\alpha(k)} \|^2 - x^* \|^2 \leq \| F(x_\beta) - y_{\delta_k} \|^2 + \alpha(\delta_k) \| x_\beta - x^* \|^2.
\]
(3.3)

This implies the boundedness of \( x_{\alpha(k)} \) and \( F(x_{\alpha(k)}) \). Therefore there exist \( \bar{x} \in X \) and \( \bar{y} \in Y \) and a subsequence of \( x_{\alpha(k)} \), for simplicity we still denote it by \( x_{\alpha(k)} \), such that \( x_{\alpha(k)} \to \bar{x} \) and \( F(x_{\alpha(k)}) \to \bar{y} \) as \( k \to \infty \), where \( \to \) denotes the weak convergence. Hence by the weakly closedness of \( F \) we have \( \bar{x} \in D(F) \) and \( F(\bar{x}) = \bar{y} \). From the weak lower semicontinuity of the Hilbert space norm it follows that
\[
\| F(\bar{x}) - y_0 \| \leq \liminf_{k \to \infty} \| F(x_{\alpha(k)}) - y_0 \| \quad \| \bar{x} - x^* \| \leq \liminf_{k \to \infty} \| x_{\alpha(k)} - x^* \|.
\]
(3.4)

This together with (3.3) gives
\[
\| F(\bar{x}) - y_0 \|^2 + \bar{\beta} \| \bar{x} - x^* \|^2 \leq \liminf_{k \to \infty} \| F(x_{\alpha(k)}) - y_{\delta_k} \|^2 + \alpha(\delta_k) \| x_{\alpha(k)} \|^2 - x^* \|^2
\]
\[
\leq \limsup_{k \to \infty} \| F(x_{\alpha(k)}) - y_{\delta_k} \|^2 + \alpha(\delta_k) \| x_{\alpha(k)} \|^2 - x^* \|^2
\]
\[
\leq \| F(x_\beta) - y_0 \|^2 + \bar{\beta} \| x_\beta - x^* \|^2
\]
which implies that \( \bar{x} \) is also a solution of the minimization problems (1.3) with \( y_\beta \) and \( \alpha \) replaced by \( y_0 \) and \( \beta \), respectively, and
\[
\lim_{k \to \infty} \| F(x_{\alpha(k)}) - y_0 \|^2 + \beta \| x_{\alpha(k)} \|^2 - x^* \|^2 = \| F(\bar{x}) - y_0 \|^2 + \beta \| \bar{x} - x^* \|^2.
\]
Combining this with (3.4) it is easy to show
\[
\| F(\bar{x}) - y_0 \| = \lim_{k \to \infty} \| F(x_{\alpha(k)}) - y_0 \| \quad \| \bar{x} - x^* \| = \lim_{k \to \infty} \| x_{\alpha(k)} - x^* \|.
\]
Since \( x_{\alpha(k)} \to \bar{x} \) and \( F(x_{\alpha(k)}) \to \bar{y} \), we have \( x_{\alpha(k)} \to \bar{x} \) and \( F(x_{\alpha(k)}) \to \bar{y} \). Now using the definition of \( \alpha(\delta_k) \) and (3.1) we can prove
\[
\beta(F(\bar{x}) - y_0, (\beta I + F'(\bar{x})F'(\bar{x}))^{-1}(F(\bar{x}) - y_0) = 0
\]
which gives \( F(\bar{x}) = y_0 \), i.e. \( \bar{x} \) is a solution of (1.1). Since the definition of \( \bar{x} \) implies \( \| \bar{x} - x^* \| \leq \| x_0 - x^* \| \), from the uniqueness of \( x^* \)-MNS it follows \( \bar{x} = x_0 \), and hence \( x_{\alpha(k)} \to x_0 \).

Finally we suppose there is a sequence \( \delta_k \) satisfying \( \delta_k \to 0 \) such that \( \alpha(\delta_k) \to 0 \) as \( k \to \infty \). Now we have \( F(x^*) \neq y_0 \) and \( \| F(x^*) - y_0 \| > c \delta_k \) for sufficiently large \( k \). Therefore the definition of \( \alpha(\delta_k) \) gives \( \| F(x_{\alpha(k)}) - y_0 \|^2 > c \delta_k^2 \). According to the definition of \( x_{\alpha(k)} \), it follows that
\[
c \delta_k^2 + \alpha(\delta_k) \| x_{\alpha(k)} \|^2 - x^* \|^2 \leq \| F(x_{\alpha(k)}) - y_0 \|^2 + \alpha(\delta_k) \| x_{\alpha(k)} \|^2 - x^* \|^2 \]
\[
\leq \delta_k^2 + \alpha(\delta_k) \| x_0 - x^* \|^2.
\]
Since \( c > 1 \) we have \( \| x_{\alpha(k)} \|^2 - x^* \|^2 \leq \| x_0 - x^* \| \) and \( \lim_{k \to \infty} F(x_{\alpha(k)}) = y_0 \). Now we can use the standard technique (cf [13]) to show \( x_{\alpha(k)} \to x_0 \) again.

As a byproduct of the argument in the proof of theorem 3.1 we have the following

**Lemma 3.1.** Under the assumptions in theorem 3.1, if \( F(x^*) \neq y_0 \) and \( x_0 \) satisfies (2.1) then
\[
\lim_{k \to \infty} \alpha(\delta_k) = 0.
\]
Proof. In fact from the proof of theorem 3.1 we see $\alpha(\delta)$ cannot have the cluster $\infty$ as $\delta \to 0$. If $\alpha(\delta)$ has a cluster $0 < \beta < \infty$, then $F(x_\beta) = y_0$, where $x_\beta$ is a solution of the minimization problem (1.3) with $y_3$ and $\alpha$ replaced by $y_0$ and $\beta$. Since $\|x_\beta - x^*\| \leq \|x_0 - x^*\|$, it follows from (2.1) that $x_\beta$ is an interior point of $D(F)$ and there holds the first-order necessary optimality condition for $x_\beta$

$$F'(x_\beta)^* (F(x_\beta) - y_0) + \beta(x_\beta - x^*) = 0.$$  

This implies $x_\beta = x^*$ and thus $F(x^*) = y_0$ which is a contradiction. □

Now we return to (3.2). It seems reasonable to make such an assumption if the problem has practical interest and $x^*$ is sufficiently close to $x_0$. However it is helpful to give some sufficient conditions. Below we will show the validity of (3.2) under the condition that

$$\|F(x) - F(z) - F'(z)(x - z)\| \leq \eta \|x - z\|, \quad x, z \in B_p(x_0)$$

(3.5) with a constant $\eta \geq 0$.

Lemma 3.2. Let (2.1) and (3.5) hold, $c \geq 2$ and $2\eta\|x_0 - x^*\| < 1$, then (3.2) is valid and $x_0$ is the unique $x^*$-MNS of (1.1).

Proof. Suppose $\alpha \geq \alpha_0$ and assume the minimization problem (1.3) has two solutions $x_\delta^\alpha$ and $\tilde{x}_\delta^\alpha$. Then

$$\|F(x_\delta^\alpha) - F(\tilde{x}_\delta^\alpha)\|^2 + \alpha \|x_\delta^\alpha - \tilde{x}_\delta^\alpha\|^2$$

$$= 2(F(\tilde{x}_\delta^\alpha) - F(x_\delta^\alpha), y_\delta - F(x_\delta^\alpha)) + 2\alpha(\tilde{x}_\delta^\alpha - x_\delta^\alpha, x^* - x_\delta^\alpha).$$

(3.6)

Since $c \geq 2$, the definitions of $x_\delta^\alpha$ and $\alpha_0$ give $\|x_\delta^\alpha - x^*\| \leq (\delta/\sqrt{\alpha}) + \|x_0 - x^*\| \leq (\delta/\sqrt{\alpha}) + \|x_0 - x^*\| \leq 2\|x_0 - x^*\|$. Therefore from (2.1) it follows that $x_\delta^\alpha$ is an interior point of $D(F)$ and there holds the first-order necessary optimality condition

$$F'(x_\delta^\alpha)^* (F(x_\delta^\alpha) - y_\delta) + \alpha(\tilde{x}_\delta^\alpha - x^*) = 0.$$  

(3.7)

Substituting (3.7) into (3.6) and using (3.5) yields

$$\|F(x_\delta^\alpha) - F(\tilde{x}_\delta^\alpha)\|^2 + \alpha \|x_\delta^\alpha - \tilde{x}_\delta^\alpha\|^2 = 2(F(\tilde{x}_\delta^\alpha) - F(x_\delta^\alpha), y_\delta - F(x_\delta^\alpha))$$

$$\leq 2\eta \|\tilde{x}_\delta^\alpha - x_\delta^\alpha\| \|F(\tilde{x}_\delta^\alpha) - F(x_\delta^\alpha)\| \|y_\delta - F(x_\delta^\alpha)\|$$

$$\leq \|F(x_\delta^\alpha) - F(\tilde{x}_\delta^\alpha)\|^2 + 2\eta \|y_\delta - F(x_\delta^\alpha)\|^2 \|x_\delta^\alpha - \tilde{x}_\delta^\alpha\|^2.$$  

(3.8)

This implies $x_\delta^\alpha = \tilde{x}_\delta^\alpha$ if we can prove $\eta \|F(x_\delta^\alpha) - y_\delta\|/\sqrt{\alpha} < 1$. By using $2\eta\|x_0 - x^*\| < 1$, from the definition of $x_\delta^\alpha$ it follows

$$\frac{\eta \|F(x_\delta^\alpha) - y_\delta\|}{\sqrt{\alpha}} \leq \eta \left( \frac{\delta}{\sqrt{\alpha}} + \|x_0 - x^*\| \right) \leq 2\eta\|x_0 - x^*\| < 1$$

which completes the proof of (3.2).

The uniqueness of $x^*$-MNS of (1.1) follows from [7, lemma 3.6]. □

4. Rate of convergence

In this section we always assume that rule 2.1 is well defined and do not state the conditions explicitly. We also use $x_\alpha$ to denote a solution of the minimization problem (1.3) with $y_\alpha$ replaced by $y_0$. In the following we will concentrate on the derivation of a suitable rate of convergence under certain conditions. A frequently used assumption is the Lipschitz continuity of the Fréchet derivative of $F$, i.e.

$$\|F'(x) - F'(z)\| \leq L \|x - z\| \quad \forall x, z \in B_p(x_0)$$

(4.1)
with some constant $L \geq 0$. As a consequence of (4.1) we have

$$\|F(x) - F(z) - F'(z)(x - z)\| \leq \frac{1}{2}L\|x - z\|^2, \quad \forall x, z \in B_p(x_0).$$

(4.2)

Since (4.1) is rather weak and provides insufficient information on $F$, we cannot establish the rates as in theorem 2.1 for each $0 < v < 2$. However when $x_0 - x^*$ is sufficiently smooth, we can give a sound result.

**Theorem 4.1.** Let (2.1) and (4.1) hold, $c > 9$, $\|F(x^*) - y_0\|^2 > c\delta^2$, and assume that there is an $\omega \in \mathcal{N}(F'(x_0)^*) \subset X$ such that $x_0 - x^* = F'(x_0)^* F'(x_0)\omega$. If $L\|u\|$ is sufficiently small, then there is a constant $C_0$ independent of $\delta$ such that

$$\|x^{\delta}_u - x_0\| \leq C_0 \delta^{2/3}$$

(4.3)

for the $x^{\delta}_u(\delta)$ determined by rule 2.1, where $u \in \mathcal{N}(F'(x_0)^*) \subset Y$ is such that $x_0 - x^* = F'(x_0)^* u$.

To prove this assertion, we note that when $2L\|u\| < 1$ there holds the stability estimate (cf [11])

$$\|x^{\delta}_u - x_0\| \leq \frac{2}{\sqrt{1 - 2L\|u\|}} \sqrt{\alpha}$$

(4.4)

for all $\alpha \geq \alpha_0$. Therefore from the triangle inequality it follows for $L\|u\|$ sufficiently small there holds

$$\|x^{\delta}_u - x_0\| \leq C_1 \frac{\delta}{\sqrt{\alpha}} + \|x_\alpha - x_0\|$$

(4.5)

with a constant $C_1$ independent of $\delta$. If we can give the estimate of $\|x_\alpha - x_0\|$ and the upper and lower bounds for $\alpha(\delta)$, then (4.3) can be proved easily. We do this according to the following lines.

**Lemma 4.1.** Let (2.1) and (4.1) hold, $c > 9$ and $\|F(x^*) - y_0\|^2 > c\delta^2$. If there is a $u \in \mathcal{N}(F'(x_0)^*) \subset Y$ such that $x_0 - x^* = F'(x_0)^* u$ and $2L\|u\| < 1$, then for the $\alpha(\delta)$ determined by rule 2.1 there holds

$$\alpha(\delta) \geq \frac{\sqrt{c} - 3\delta}{2\|u\|}.$$ 

(4.6)

**Proof.** Since $\|F(x^*) - y_0\|^2 > c\delta^2$, the definition of $\alpha(\delta)$ implies $\|F(x^{\delta}_u(\delta)) - y_0\| \geq \sqrt{c}\delta$. Since the proof of [11, theorem 1] implies

$$\|F(x^{\delta}_u) - F(x_\alpha)\| \leq 2\delta$$

(4.7)

for all $\alpha \geq \alpha_0$, we have from (1.2) that $\|F(x^{\delta}_u(\delta)) - y_0\| \geq (\sqrt{c} - 3)^2\delta$. Following the proof of theorem 2.4 in [2] it gives for all $\alpha > 0$

$$\|F(x_\alpha) - y_0\| \leq 2\|u\|\alpha.$$ 

(4.8)

Therefore $2\|u\|\alpha(\delta) \geq (\sqrt{c} - 3)^2\delta$ and the proof follows. □

**Lemma 4.2.** Under the assumptions in lemma 4.1, if $L\|u\|$ is sufficiently small such that

$$q_0 := \left(\frac{8}{\sqrt{1 - 2L\|u\|} (\sqrt{c} - 3)} + \frac{4}{\sqrt{1 - L\|u\|}}\right) L\|u\| \leq 1$$

(4.9)

then for the $\alpha := \alpha(\delta)$ defined by rule 2.1 there holds

$$\frac{(\sqrt{c} - 3)^2}{3} \delta \leq \alpha(F(x_\alpha) - y_0, (\alpha I + B_0)^{-1}(F(x_\alpha) - y_0)) \leq 3(\sqrt{c} + 3)^2\delta^2$$

(4.10)

where $B_0 := F'(x_0) F'(x_0)^*$. 

Tikhonov regularization of nonlinear ill-posed problems

1093
Proof. Assume $B_0^3 := F'(x_0^3)F'(x_0^3)^*$, then from (1.2) and (4.7) it immediately gives
$$(\sqrt{\alpha} - 3)^{2} \leq (\alpha(F(x_0) - y_0), (\alpha I + B_0^3)^{-1}(F(x_0) - y_0)) \leq (\sqrt{\alpha} + 3)^{2} \delta^2.$$ 

Let
$$\begin{align*}
a &:= \alpha(F(x_0) - y_0), (\alpha I + B_0^3)^{-1}(F(x_0) - y_0)), \\
b &:= \alpha(F(x_0) - y_0), (\alpha I + B_0)^{-1}(F(x_0) - y_0)), \\
F &= (\alpha I + B_0^3)^{-1/2}(B_0 - B_0^3)(\alpha I + B_0)^{-1/2}
\end{align*}$$

then we have
$$|a - b| = |\alpha(F(x_0) - y_0), (\alpha I + B_0^3)^{-1}(B_0 - B_0^3)(\alpha I + B_0)^{-1}(F(x_0) - y_0))|$$

$$= |\alpha((\alpha I + B_0^3)^{-1/2}(F(x_0) - y_0), F(\alpha I + B_0)^{-1/2}(F(x_0) - y_0))|$$

$$\leq \frac{2}{\sqrt{\alpha}} \|F'(x_0^3) - F'(x_0)\| \leq \frac{2L}{\sqrt{\alpha}} \|x_0^3 - x_0\|.$$ 

Following the proof of [2, theorem 2.4] we also have

$$\|x_0 - x_0\| \leq \frac{2\sqrt{\alpha}}{\sqrt{1 - L\|u\|}}.$$  

(4.11)

Therefore from (4.4) and (4.6) we obtain

$$\|\mathcal{F}\| \leq 2L \left( \frac{2}{\sqrt{1 - 2L\|u\|}} \frac{\delta}{\alpha} + \frac{2\|u\|}{\sqrt{1 - L\|u\|}} \right) \leq q_0 \leq 1.$$ 

Thus $|a - b| \leq \frac{1}{\alpha} (a + b)$ which implies $a/3 \leq b \leq 3a$ and the assertion follows.  

Now we give the estimate of $\|x_0 - x_0\|$ and the upper and lower bounds for $\alpha(\delta)$.

Lemma 4.3. Let (2.1) and (4.1) hold and assume $x_0 - x^* = (F'(x_0)^*F'(x_0))^{1/2} \omega$ with some $1 < \nu < 2$ and $\omega \in N(F'(x_0)^*) \subset X$. If $u \in N(F'(x_0)^*) \subset Y$ denotes the element such that $x_0 - x^* = F'(x_0)^*u$ and if $L\|u\|$ is sufficiently small such that

$$q_1 := \left( 4 + \frac{1}{\sqrt{1 - L\|u\|}} \right) \frac{\|u\|}{L\|u\|} \leq 1$$  

(4.12)

then for each $\alpha > 0$ there holds

$$\|x_0 - x_0\| \leq 2\alpha^{\nu/2}\|\omega\|.$$  

(4.13)

Proof. Since $x_0$ is an interior point of $D(F)$, we have

$$F'(x_0)^*(F(x_0) - y_0) + \alpha(x_0 - x^*) = 0.$$  

(4.14)

By letting $A_0 := F'(x_0)^*F'(x_0)$ and

$$\begin{align*}
\hat{x}_o &:= x_0 + \alpha(\alpha I + A_0)^{-1}(x^* - x_0), \\
r_o &:= (F'(x_0)+F'(x_0)^*)(F(x_0) - y_0), \\
s_o &:= F'(x_0)^*(F(x_0) - F(x_0) - F'(x_0)(x_0 - x_0))
\end{align*}$$

we have 

$$s_o \leq (\hat{x}_o - x_0)^2 \leq \frac{2}{\sqrt{\alpha}} \|x_0^3 - x_0\|.$$

By letting $\hat{A}_0 := F'(x_0)^*F'(x_0)$ and 

$$\begin{align*}
\hat{x}_o &:= x_0 + \alpha(\alpha I + \hat{A}_0)^{-1}(x^* - a), \\
r_o &:= (\hat{A}_o + (\hat{A}_o)^*)/2, \\
\hat{s}_o &:= (\hat{A}_o + (\hat{A}_o)^*)/2
\end{align*}$$

we have 

$$\hat{s}_o \leq (\hat{x}_o - x_0)^2 \leq \frac{2}{\sqrt{\alpha}} \|x_0^3 - x_0\|.$$
it follows from (4.14) that

\[ x_\alpha = \hat{x}_\alpha - (\alpha I + A_0)^{-1} (r_\alpha + s_\alpha). \tag{4.15} \]

Therefore from (4.1) and (4.2) it follows

\[
\| x_\alpha - \hat{x}_\alpha \| = \left\| (\alpha I + A_0)^{-1} (r_\alpha + s_\alpha) \right\|
\leq \| r_\alpha \| + \frac{1}{2\sqrt{\alpha}} \| F(x_\alpha) - y_0 - F'(x_0)(x_\alpha - x_0) \|
\leq \frac{1}{\alpha} \| F'(x_\alpha) - F'(x_0) \| \| F(x_\alpha) - y_0 \| + \frac{L\| x_\alpha - x_0 \|^2}{4\sqrt{\alpha}}
\leq \left( \frac{L\| F(x_\alpha) - y_0 \|}{\alpha} + \frac{L\| x_\alpha - x_0 \|}{4\sqrt{\alpha}} \right) \| x_\alpha - x_0 \|.
\]

Using (4.8) and (4.11) we obtain

\[
\| x_\alpha - \hat{x}_\alpha \| \leq \frac{1}{2q_1} \| x_\alpha - x_0 \| \leq \frac{1}{2} \| x_\alpha - x_0 \|, \text{ which gives }
\| x_\alpha - x_0 \| \leq 2\| \hat{x}_\alpha - x_0 \| = 2\| \alpha(\alpha I + A_0)^{-1} A_0^{1/2} \alpha \| \leq 2^{1/2} \| \alpha \|.
\]

\[ \square \]

**Lemma 4.4.** Under the assumptions in lemma 4.3, if \( L\| u \| \) is sufficiently small such that (4.9) and (4.12) hold, then

\[
\alpha(\delta) \geq \left( \frac{\sqrt{c} - 3}{2\sqrt{3}(\sqrt{1 - L\| u \|} + L\| u \|)} \| \alpha \| \right)^{2/(1+\nu)} \tag{4.16}
\]

for the \( \alpha := \alpha(\delta) \) determined by rule 2.1.

**Proof.** From (4.2), (4.11) and (4.13) we have

\[
\sqrt{\alpha} \| (\alpha I + B_0)^{-1/2} (F(x_\alpha) - y_0) \| \leq \sqrt{\alpha} \| x_\alpha - x_0 \| + \frac{1}{2} L\| x_\alpha - x_0 \|^2
\leq \sqrt{\alpha} \left( 1 + \frac{L\| u \|}{\sqrt{1 - L\| u \|}} \right) \| x_\alpha - x_0 \|
\leq 2 \left( 1 + \frac{L\| u \|}{\sqrt{1 - L\| u \|}} \right) \| \alpha \| \alpha^{(1+\nu)/2}.
\]

Therefore it follows from (4.10) that

\[
\frac{\sqrt{c} - 3}{\sqrt{3}} \delta \leq 2 \left( 1 + \frac{L\| u \|}{\sqrt{1 - L\| u \|}} \right) \| \alpha \| \alpha^{(1+\nu)/2}
\]

which gives (4.16). \[ \square \]

**Lemma 4.5.** Under the assumptions in lemma 4.4, there holds

\[
\liminf_{\delta \to 0} \frac{\delta^{2/3}}{\alpha(\delta)} > 0. \tag{4.17}
\]

**Proof.** Let

\[ I(\delta) := \frac{\| (\alpha(\delta) I + B_0)^{-1/2} F'(x_0)(x_{\alpha(\delta)} - x_0) \|}{\alpha(\delta)} \]

then

\[ I(\delta) \geq \frac{\| (\alpha(\delta) I + B_0)^{-1/2} F'(x_0)(\hat{x}_{\alpha(\delta)} - x_0) \|}{\alpha(\delta)} = \frac{\| x_{\alpha(\delta)} - \hat{x}_{\alpha(\delta)} \|}{\alpha(\delta)} = I_1(\delta) - I_2(\delta). \]
If we use \( \{E_i\} \) to denote the spectral family generated by \( A_0 \), then from the smoothness assumption on \( x_0 - x^* \) we have

\[
I_1(\delta) = \frac{1}{\alpha(\delta)} ((\alpha(\delta) I + A_0)^{-1} A_0(\hat{x}_{\alpha(\delta)} - x_0), \hat{x}_{\alpha(\delta)} - x_0)
\]

\[
= ((\alpha(\delta) I + A_0)^{-3} A_0^3 \omega, \omega)
\]

\[
= \int_0^{\infty} \frac{\lambda^3}{(\alpha(\delta) + \lambda)^3} d(E_\omega \omega, \omega).
\]

Since \( \omega \in \mathcal{N}(F'(x_0))^{-1} \) and \( \alpha(\delta) \to 0 \) implied by lemma 3.1, by taking \( \delta \to 0 \) we have \( I_1(\delta) \to \|\omega\| \). For \( I_2(\delta) \) we can use (4.15), (4.1), (4.2), (4.13) and (4.8) to obtain

\[
I_2(\delta) \leq \frac{\|x_{\alpha(\delta)}\| + \|\alpha(\delta) I + B_0\|^{-1} x_{\alpha(\delta)}\|}{\alpha(\delta)^2}
\]

\[
\leq L \|x_{\alpha(\delta)} - x_0\| \|F(x_{\alpha(\delta)}) - y_0\| + L \|x_{\alpha(\delta)} - x_0\|^2 \frac{\|\alpha(\delta)\|^2}{\alpha(\delta)^3/2}
\]

\[
\leq 4L \|u\|^{1/2} \alpha(\delta)^{1/2} \to 0 \quad \text{as} \quad \delta \to 0.
\]

Therefore

\[
\liminf_{\delta \to 0} \frac{\|\alpha(\delta) I + B_0\|^{-1/2} (F(x_{\alpha(\delta)}) - y_0)\|}{\alpha(\delta)} \geq (1 - 4L \|u\|) \|\omega\| > 0.
\]

By using (4.10) we obtain (4.17) immediately.

Let us return to the proof of theorem 4.1. Clearly, lemma 4.5 implies \( \alpha(\delta) \leq C_2 \delta^{2/3} \) with a constant \( C_2 \) independent of \( \delta \). Combining this with (4.16), (4.13) and (4.5), assertion (4.3) follows immediately and the proof of theorem 4.1 is complete.

Now we consider the question whether we can obtain the rate of convergence under the assumption \( x_0 - x^* \in R((F'(x_0)^{-1} F'(x_0))^{1/2}) \) with \( \nu < 2 \). By checking the proof of theorem 4.1, it is easy to see that the big burden is to get the estimate like \( \alpha(\delta) \leq C_2 \delta^{2/(1+\nu)} \) with some constant \( C \) independent of \( \delta \). Such an estimate for \( \nu = 2 \) is given in lemma 4.5 which takes account of the saturation of Tikhonov regularization to obtain \( I_1(\delta) \to \|\omega\| \) as \( \delta \to 0 \). It seems there is no hope to obtain such an estimate for \( \nu < 2 \); this fact can be clarified from the discussion in [9]. Thus, in order to derive the expected rates other techniques should be explored.

We conclude this section by a remark that rule 2.1 is well defined under the conditions in lemma 4.1. In fact, we need only to verify (3.2). Let \( \alpha \geq \alpha_0 \) and let \( x_0^1 \) and \( \hat{x}_0^1 \) be the two solutions of (1.3), it follows from (3.8) and (4.2) that \( \alpha \|x_0^1 - \hat{x}_0^1\|^2 \leq L \|F(x_0^1) - y_0\| \|x_0^1 - \hat{x}_0^1\|^2 \). Since (4.7), (4.8) and (4.6) imply

\[
\frac{L \|F(x_0^1) - y_0\|}{\alpha} \leq L \left( \frac{\delta}{\alpha} + 2 \|u\| \right) \leq \frac{3L \|x_0 - x^*\|}{\sqrt{c - 1}} + 2L \|u\|
\]

we have \( \hat{x}_0^1 = x_0^1 \) provided that \( L \|u\| \) is sufficiently small. Of course, this remark has more interest in theoretical analysis rather than in practical applications.
5. Applications

The results presented in sections 3 and 4 can be applied to many concrete problems. Below we give only a few such examples to illustrate the conditions required in the foregoing sections.

Example 5.1. We continue the discussion of example 2.1 with \(g_0 = g_1 = 0\). Now the Fréchet derivative is given by

\[ F'(a)h = A(a)^{-1}(hu_t(a)), \]

where \(A(a) : H^2 \cap H^1_0 \mapsto L^2\) is given by \(A(a)u = -(au_t)_t\), which is an isomorphism uniformly in a neighbourhood of \(a_0\). Obviously (3.1) and (4.1) are trivial. From [6] we know (3.2) is also true. We note that the verification of (3.5) is given in [1]. Hence theorems 3.1 and 4.1 are applicable.

Example 5.2. Considering the problem of estimating \(c\) in

\[
\begin{aligned}
-\Delta u + cu &= f \\
u &= 0
\end{aligned}
\]

in \(\Omega\) on \(\partial \Omega\) (5.1)

where \(\Omega\) is a bounded domain in \(R^2\) or \(R^3\) with smooth boundary or with \(\Omega\) being a parallelepiped and \(f \in L^2(\Omega)\).

The nonlinear operator \(F : D(F) \subset L^2(\Omega) \mapsto L^2(\Omega)\) is defined as the parameter-to-solution mapping

\[ F(c) = u(c) \]

with \(u(c)\) the unique solution of (5.1). \(F\) is well defined on

\[ D(F) := \{c \in L^2 : \|c - \hat{c}\|_{L^2} \leq \gamma \text{ for some } \hat{c} \geq 0 \text{ a.e.}\} \]

with some \(\gamma > 0\), cf [12]. It is easy to show that \(F\) is Fréchet differentiable and the Fréchet derivative is given by

\[ F'(c)h = -A(c)^{-1}(hu(c)) \]

where \(A(c) : H^2 \cap H^1_0 \mapsto L^2\) defined by \(A(c)u = -\Delta u + cu\) is an isomorphism uniformly in a neighbourhood \(U(c_0)\) of the sought solution \(c_0\). From the discussion in [12] we know (2.2) and (2.3) are not true for this example. Let us check the assumptions of this paper. Obviously, (3.1) is trivial and (4.1) is an easy exercise. In order to guarantee (3.2), let us verify (3.5). Suppose \(c, d \in U(c_0)\), then it is easy to show

\[ A(c)(F(d) - F(c)) = (c - d)F(d). \]

Therefore

\[ \|F(d) - F(c) - F'(c)(d - c)\|_{L^2} = \|A(c)^{-1}((c - d)(F(c) - F(d)))\|_{L^2} \leq \eta\|c - d\|_{L^2}\|F(c) - F(d)\|_{L^2} \]

with some constant \(\eta\), and (3.5) holds locally.

Example 5.3. We treat the ill-posed nonlinear integral equation of autoconvolution type defined on the interval \([0, 1]\)

\[ F(x)(s) := \int_0^s x(s - t)x(t) \, dt = y(s) \]
The nonlinear operator \( F : L^2[0, 1] \rightarrow L^2[0, 1] \) is Fréchet differentiable with derivative
\[
(F'(x)h)(s) = 2 \int_0^s x(s - t)h(t) \, dt.
\]

Obviously, (3.1) is valid. Although (3.5) is unsatisfactory everywhere, we can verify (4.1) as follows:
\[
\|F'(x) - F'(z)\|_{L^2} = 2 \left\{ \int_0^1 \left[ \int_0^s (x(s - t) - z(s - t))h(t) \, dt \right]^2 \, ds \right\}^{1/2} \\
\leq 2 \left\{ \int_0^1 \int_0^s |x(s - t) - z(s - t)|^2 \, dt \int_0^s h(t)^2 \, dt \, ds \right\}^{1/2} \\
\leq 2 \|x - z\|_{L^2} \|h\|_{L^2}, \quad \forall h \in L^2[0, 1].
\]

Therefore (3.2) can be guaranteed if the sought solution \( x_0 \) is sufficiently smooth. However, this statement has little use in practice since it is not easy to obtain the information on \( L^2 \|u\| \).

We conjecture that it is possible to prove (3.2) directly for this example if \( x^* \) is sufficiently close to the true solution \( x_0 \).

**Acknowledgments**

I would like to thank Professor C W Groetsch of the University of Cincinnati for sending me the monograph [5] which provided me with more background about ill-posed problems. I am also indebted to Dr O Scherzer of the Universität of Linz, Austria for his kindly providing me with the offprint of [11]. Thanks also go to the referees for their careful reading of the manuscript. This work is supported by the National Natural Science Foundation of China.

**References**