Inexact Newton–Landweber iteration for solving nonlinear inverse problems in Banach spaces

Qinian Jin
Mathematical Sciences Institute, The Australian National University, Canberra, ACT 0200, Australia
E-mail: Qinian.Jin@anu.edu.au

Received 14 June 2011, in final form 9 January 2012
Published 24 April 2012
Online at stacks.iop.org/IP/28/065002

Abstract
By making use of duality mappings, we formulate an inexact Newton–Landweber iteration method for solving nonlinear inverse problems in Banach spaces. The method consists of two components: an outer Newton iteration and an inner scheme providing the increments by applying the Landweber iteration in Banach spaces to the local linearized equations. It has the advantage of reducing computational work by computing more cheap steps in each inner scheme. We first prove a convergence result for the exact data case. When the data are given approximately, we terminate the method by a discrepancy principle and obtain a weak convergence result. Finally, we test the method by reporting some numerical simulations concerning the sparsity recovery and the noisy data containing outliers.

(Some figures may appear in colour only in the online journal)

1. Introduction
We are interested in nonlinear inverse problems which can be formulated as

\[ F(x) = y, \]  \hspace{1cm} (1.1)

where \( F : \mathcal{D}(F) \subset \mathcal{X} \rightarrow \mathcal{Y} \) is a nonlinear operator between two Banach spaces \( \mathcal{X} \) and \( \mathcal{Y} \). We assume that (1.1) has a solution \( x^* \) in the domain \( \mathcal{D}(F) \) of \( F \) and that \( F \) is Fréchet differentiable with the Fréchet derivative \( F'(x) : \mathcal{X} \rightarrow \mathcal{Y} \). Assuming \( \mathcal{X} \) and \( \mathcal{Y} \) are Hilbert spaces, many regularization methods have been developed for solving nonlinear inverse problems in the last two decades; see [1, 3, 4, 6–9, 12, 13, 15] and references therein. Regularization methods in Hilbert spaces can produce good results when the sought solution is smooth and the data contain only Gaussian noise. However, because such methods have a tendency to oversmooth solutions and are susceptible to the types of noise, they may not produce good results in applications where the sought solution is sparse, or contains discontinuities, or the data contain large noise points due to procedural error. Instead, regularization in Banach spaces can be used...
to obtain better reconstruction results. Moreover, due to their intrinsic feature, many inverse problems are more naturally formulated in Banach spaces than in Hilbert spaces. Therefore, it is necessary to develop methods to solve nonlinear inverse problems in Banach spaces.

Due to the variational formulation, Tikhonov regularization can be easily adapted to solve nonlinear inverse problems in Banach spaces and some convergence analysis has been carried out in recent years; see [14] and references therein. Since the numerical realization requires one to solve several minimization problems which are non-convex and hence their minimizers are not necessarily unique, Tikhonov regularization in general is rather expensive. It is still a challenging topic to find the minimizers of nonlinear Tikhonov functionals efficiently.

The iteratively regularized Gauss–Newton method has been extended in [11] to solve nonlinear inverse problems in Banach spaces and the convergence rates have been derived in [10] under suitable source conditions. In this method, each iteration requires solving a convex minimization problem in Banach space. In many important applications, the convex functionals involved in these minimization problems are not smooth enough and thus the well-known efficient solvers cannot be applied directly. How to find fast and efficient methods for solving these convex minimization problems is indeed an emerging issue.

By making use of duality mappings and the idea in [16] of generalizing linear Landweber iteration to solve linear inverse problems in Banach spaces, the nonlinear Landweber method in [4] has been extended in [11] to the Banach space setting. This is a fully iterative method and its implementation is straightforward. However, this method is known to be slow convergent.

Inspired by the inexact Newton regularization methods in Hilbert spaces developed in [3, 15, 13] successively, in this paper we will propose an inexact Newton–Landweber iteration method for solving nonlinear inverse problems in Banach spaces. The method consists of an outer Newton iteration and an inner scheme providing increments by regularizing the local linearized equations via the Landweber iteration in [16]. To be more precise, we start with an initial guess $x_0$. If $x_n$ is a current iterate, we consider the linearized equation

$$F'(x_n)(x - x_n) = y - F(x_n)$$

and apply the nonlinear version of the linear Landweber iteration in [16] to produce a sequence of approximate solutions which enables us to define the next iterate $x_{n+1}$ such that

$$\|y - F(x_n) - F'(x_n)(x_{n+1} - x_n)\| \leq \mu \|y - F(x_n)\|$$

for some preassigned number $0 < \mu < 1$. In case the data are given approximately, a discrepancy principle is used to terminate the iteration. Comparing with the nonlinear Landweber iteration in Banach spaces formulated in [11], our method has the advantage of reducing computational work by computing more cheap steps in each inner scheme. It is worthwhile pointing out that there have been some attempts to accelerate the Landweber iteration for inverse problems in Banach spaces [17, 5]. We expect that our method can become more favorable by using these accelerated versions in the inner scheme. We hope to address this issue in a forthcoming paper.

This paper is organized as follows. In section 2, we briefly review some basic geometric properties of Banach spaces. In section 3, by making use of duality mappings, we first formulate an inexact Newton–Landweber iteration method in Banach spaces and show that it is convergent for the exact data case. When the data are given approximately, we terminate the method by a discrepancy principle and show that the method is weakly convergent. Finally, in section 4, we test the method by reporting some numerical results concerning the sparsity recovery and the noisy data containing outliers.
2. Preliminaries

Let $\mathcal{X}$ and $\mathcal{Y}$ be two Banach spaces whose norms are denoted by $\| \cdot \|$. We use $\mathcal{X}^*$ and $\mathcal{Y}^*$ to denote their dual spaces, respectively. Given $x \in \mathcal{X}$ and $x^* \in \mathcal{X}^*$ we write $\langle x^*, x \rangle = x^*(x)$ for the duality pair. We use $L(\mathcal{X}, \mathcal{Y})$ to denote the collection of all continuous linear operators from $\mathcal{X}$ to $\mathcal{Y}$. For $T \in L(\mathcal{X}, \mathcal{Y})$, we use $T^* : \mathcal{Y}^* \rightarrow \mathcal{X}^*$ to denote its dual, i.e. $(T^* y^*) = \langle y^*, T x \rangle$ for any $x \in \mathcal{X}$ and $y^* \in \mathcal{Y}^*$. The dual operator $T^*$ is in $L(\mathcal{Y}^*, \mathcal{X}^*)$ with $\|T^*\| = \|T\|$. For any number $1 < p < \infty$, we use $\rho^p$ to denote the number conjugate to $p$, i.e. $1/p + 1/p^* = 1$.

In order to formulate our method, we first review some geometric properties of Banach spaces which can be found in [2]. For a Banach space $\mathcal{X}$, we may introduce the modulus of smoothness

$$\rho_X(\tau) := \sup\{\|x + \bar{x}\| + \|x - \bar{x}\| - 2 : \|x\| = 1, \|\bar{x}\| \leq \tau, \quad \tau \geq 0\}$$

and the modulus of convexity

$$\delta_X(\epsilon) := \inf\{2 - \|x + \bar{x}\| : \|x\| = \|\bar{x}\| = 1, \|x - \bar{x}\| \geq \epsilon\}, \quad 0 \leq \epsilon \leq 2.$$ We call $\mathcal{X}$ uniformly smooth if $\lim_{\tau \to 0} \frac{\rho_X(\tau)}{\tau} = 0$ and uniformly convex if $\delta_X(\epsilon) > 0$ for all $0 < \epsilon \leq 2$. The space $\mathcal{X}$ is said to be $p$-smooth for some $p > 1$ if $\rho_X(\tau) \leq C_p \tau^p$ for all $\tau \geq 0$, while it is said to be $p$-convex with $p > 0$ if $\delta_X(\epsilon) \geq c_p \epsilon^p$ for $0 \leq \epsilon \leq 2$, where $C_p$ and $c_p$ are some positive constants. Any uniformly smooth or uniformly convex Banach spaces are reflexive. A Banach space $\mathcal{X}$ is uniformly smooth (resp. uniformly convex) if and only if $\mathcal{X}^*$ is uniformly convex (resp. uniformly smooth). $\mathcal{X}$ is $p$-convex with $1 < p < \infty$ if and only if $\mathcal{X}^*$ is $p^*$-smooth.

Given $1 < p < \infty$, the set-valued mapping $J_p : \mathcal{X} \rightarrow 2^{\mathcal{X}^*}$ defined by

$$J_p(x) = \{x^* \in \mathcal{X}^* : \|x^*\| = \|x\|^{p-1} \text{ and } \langle x^*, x \rangle = \|x\|^p\}$$

is called the duality mapping with the gauge function $t \mapsto t^{p-1}$. A selection of the duality mapping $J_p$ is defined to be a single-valued mapping $j_p : \mathcal{X} \rightarrow \mathcal{X}^*$ with the property $j_p(x) \in J_p(x)$ for each $x \in \mathcal{X}$. $J_p$ in general is multi-valued. However, if $\mathcal{X}$ is uniformly smooth, then $J_p$ is single-valued and uniformly continuous on bounded sets. If in addition $\mathcal{X}$ is uniformly convex, then $J_p$ is bijective; the inverse $J_p^{-1} : \mathcal{X}^* \rightarrow \mathcal{X}$ admits the property

$$J_p^{-1}(x^*) = J_p^* (x^*) \quad \forall x^* \in \mathcal{X}^*,$$ (2.1)

where $J_p^* : \mathcal{X}^* \rightarrow \mathcal{X}$ denotes the duality mapping with the gauge function $t \mapsto t^{p-1}$.

For a uniformly smooth Banach space $\mathcal{X}$, from the characterization of uniform smoothness in [18], it follows that for any $1 < q < \infty$ there is a positive constant $C_q$ such that

$$\frac{1}{q} \|x\|^q - \frac{1}{q} \|\bar{x}\|^q - \langle J_q(\bar{x}), x - \bar{x} \rangle \leq C_q \int_0^1 \max\{\|\bar{x}\|, \|\bar{x} + t(x - \bar{x})\|\} \frac{t\|x - \bar{x}\|}{\max\{\|\bar{x}\|, \|\bar{x} + t(x - \bar{x})\|\}} dt$$

for all $x, \bar{x} \in \mathcal{X}$. If $\mathcal{X}$ is $s$-smooth for some $s > 1$, i.e. $\rho_X(\tau) \leq C_s \tau^s$ for some constant $C_s > 0$, then we have

$$\frac{1}{q} \|x\|^q - \frac{1}{q} \|\bar{x}\|^q - \langle J_q(\bar{x}), x - \bar{x} \rangle \leq C_q C_s \|x - \bar{x}\|^s \int_0^1 t^{s-1} \max\{\|\bar{x}\|, \|\bar{x} + t(x - \bar{x})\|\}^{q/s} dt$$

which implies immediately that
1 \frac{q}{q} \|x\|^q - \frac{q}{q} \|\bar{x}\|^q - \langle J_q(\bar{x}), x - \bar{x} \rangle \leq C_{q,s} \|x - \bar{x}\|^s \left( \|\bar{x}\|^q + \|x - \bar{x}\|^{q-s} \right) \quad (2.2)

for some positive constant $C_{q,s}$ depending only on $q$ and $s$.

The most commonly used uniformly smooth and uniformly convex Banach spaces are the sequence spaces $l^p$, the Lebesgue spaces $L^p$, the Sobolev spaces $W^{k,p}$ and the Besov spaces $B^{k,p}$ with $1 < p < \infty$. It is well known that they are 2-convex and $p$-smooth if $1 < p \leq 2$ and they are $p$-convex and 2-smooth if $2 \leq p < \infty$.

In order to study the convergence property of our method, we will use the Bregman distance. When $\mathcal{X}$ is uniformly smooth, for any $1 < p < \infty$ the duality mapping $J_p$ is single valued and we can introduce the associated Bregman distance

$$\Delta_p(x, \bar{x}) := \frac{1}{p} \|x\|^p - \frac{1}{p} \|\bar{x}\|^p - \langle J_p(\bar{x}), x - \bar{x} \rangle.$$ By straightforward calculation we can see for any $x, x_1, x_2 \in \mathcal{X}$ that

$$\Delta_p(x, x_1) - \Delta_p(x, x_2) = \Delta_p(x_2, x_1) + \langle J_p(x_1) - J_p(x_2), x_2 - x \rangle. \quad (2.3)$$ Since $\|\bar{x}\|^p = \langle J_p(\bar{x}), \bar{x} \rangle$ and $1/p + 1/p^* = 1$, it is easy to see that

$$\Delta_p(x, \bar{x}) = \frac{1}{p} \|x\|^p + \frac{1}{p^*} \|\bar{x}\|^p - \langle J_p(\bar{x}), x \rangle. \quad (2.4)$$ Since $\|J_p(\bar{x})\| = \|\bar{x}\|^{p-1}$, this implies that

$$\Delta_p(x, \bar{x}) \geq \frac{1}{p} \|x\|^p + \frac{1}{p^*} \|\bar{x}\|^p - \|\bar{x}\|^{p-1} \|x\|. \quad (2.5)$$

Thus we may use Young’s inequality to conclude that $\Delta_p(x, \bar{x}) \geq 0$. If $\{x_n\} \subset \mathcal{X}$ is a sequence such that $\{\Delta_p(x_n, x)\}$ (or $\{\Delta_p(x_n, x)\}$) is bounded, then it follows easily from (2.5) that $\{x_n\}$ is a bounded sequence in $\mathcal{X}$. When $\mathcal{X}$ is also uniformly convex, by using the characterization of uniform convexity of Banach spaces in [18] it has been shown in [16] that, for a sequence $\{x_n\} \subset \mathcal{X}$, $\Delta_p(x_n, x) \to 0$ as $m, n \to \infty$ if and only if $\{x_n\}$ is a Cauchy sequence in $\mathcal{X}$.

### 3. The method

In this section, by generalizing the idea in [3, 15, 13] on the inexact Newton regularization methods in Hilbert spaces and making use of duality mappings, we will formulate an inexact Newton–Landweber iteration method for solving (1.1) in Banach spaces, where $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces with $\mathcal{X}$ being uniformly smooth and uniformly convex. We will assume that (1.1) has a solution $x^* \in D(F)$ and that $F$ is Fréchet differentiable with Fréchet derivative $F'(x) : \mathcal{X} \to \mathcal{Y}$ and the adjoint $F'(x)^* : \mathcal{Y}^* \to \mathcal{X}^*$. We will work under the following standard condition.

**Assumption 3.1.** (a) For some $\rho > 0$ there holds

$$B_p(x^*, \Delta_p) := \{x \in \mathcal{X} : \Delta_p(x, x) \leq \rho \} \subset D(F).$$

(b) There is a constant $C_0$ such that $\|F'(x)\| \leq C_0$ for all $x \in B_p(x^*, \Delta_p)$. 

(c) There is a constant $0 \leq \eta < 1$ such that

$$\|F(x) - F(\bar{x}) - F'(\bar{x})(x - \bar{x})\| \leq \eta \|F(x) - F(\bar{x})\|$$

for all $x, \bar{x} \in B_p(x^*, \Delta_p)$.

The method we will formulate consists of two components: the outer Newton iteration and the inner scheme providing the increments by applying the Landweber iteration in Banach spaces to the local linearized equations. In the formulation, we will take $1 < p < \infty$, $r \geq 1$ and
let $J_p : \mathcal{X} \to \mathcal{X}^*$ be the duality mapping with gauge function $t \mapsto t^{p-1}$ and let $j_r : \mathcal{Y} \to \mathcal{Y}^*$ be a mapping with the properties

\[
\langle j_r(z), z \rangle = \|z\|^r \quad \text{and} \quad \|j_r(z)\| \leq \|z\|^{r-1}
\]

for all $z \in \mathcal{Y}$. When $1 < r < \infty$, we may take $j_r$ to be a single-valued selection of the duality mapping on $\mathcal{Y}$ with gauge function $t \mapsto t^{r-1}$.

### 3.1. Exact data case

We first formulate the method when the data are given exactly. Let $x_0 \in \mathcal{D}(F) \subset \mathcal{X}$ be an initial guess. When $x_n$ is defined, we then define the next iterate $x_{n+1}$ by the following procedure: we first construct a sequence $\{u_{n,k}\} \subset \mathcal{X}^*$ iteratively by setting $u_{n,0} = 0$ and

\[
u_{n,k+1} = u_{n,k} + \omega_{n,k} T_n^p j_r(y - F(x_n) - T_n(z_{n,k} - x_n)),
\]

where $T_n = F'(x_n)$, $\omega_{n,k}$ is a positive number and

\[
z_{n,k} = J_p^{\omega}(J_p(x_n) + u_{n,k}).
\]

Let $k_n$ be the first integer such that

\[
\|y - F(x_n) - T_n(z_{n,k} - x_n)\| \leq \mu \|y - F(x_n)\|
\]

where $0 < \mu < 1$ is a preassigned number. We then define the next iterate to be $x_{n+1} = z_{n,k_n}$.

According to (2.1), it is easy to see that

\[
J_p(z_{n,k_n}) = J_p(x_n) + u_{n,k_n},
\]

and in particular

\[
J_p(x_{n+1}) = J_p(x_n) + u_{n,k_n}.
\]

The following result shows that $\{x_n\}$ is well defined and converges to a solution of equation (1.1).

**Theorem 3.1.** Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces with $\mathcal{X}$ being uniformly smooth and $s$-convex for some $s > 1$; let assumption 3.1 hold with $0 \leq \eta < 1/3$, $x_0 \in B_p(x^\dagger, \Delta_p)$, and $1 < p \leq s$ and $r \geq 1$ in the definition of $[x_n]$. Then the method is well defined if $\eta < \mu < 1 - 2\eta$ and $\omega_{n,k}$ is chosen as

\[
\omega_{n,k} = \min \left\{ \omega^{(1)}_{n,k}, \omega^{(2)}_{n,k} \right\},
\]

where

\[
\omega^{(1)}_{n,k} = \theta_1 \|T_n\|^{-p} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^{s-r},
\]

\[
\omega^{(2)}_{n,k} = \theta_2 \|T_n\|^{-s} \|z_{n,k}\|^{1-s} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^{s-r}
\]

for some positive constants $\theta_1$ and $\theta_2$ satisfying

\[
c_0 := 1 - \frac{\eta}{\mu} - C_{p,s} \left( \theta_1^{p-1} + \theta_2^{s-1} \right) > 0.
\]

Moreover, $\{x_n\}$ converges strongly in $\mathcal{X}$ to a solution of (1.1) in $B_p(x^\dagger, \Delta_p)$. If $x^\dagger$ is the unique solution of (1.1) in $B_p(x^\dagger, \Delta_p)$, then $x_n \to x^\dagger$ as $n \to \infty$.

**Proof.** We first prove that the method is well defined by showing that each inner iteration terminates after finite steps and that $\Delta_p(x^\dagger, x_n)$ is monotonically decreasing, i.e.

\[
\Delta_p(x^\dagger, x_{n+1}) \leq \Delta_p(x^\dagger, x_n), \quad n = 0, 1, \ldots
\]
Since \( z_{n,k} = x_{n+1} \) and \( z_{n,0} = x_n \), it suffices to show that each \( k_n \) is finite and
\[
\Delta_p(x^i, z_{n,k+1}) \leq \Delta_p(x^i, z_{n,k}), \quad k = 0, \ldots, k_n - 1.
\]
(3.2)

If \( F(x_n) = y \), then \( k_n = 0 \) and nothing needs to be proved. Therefore we may assume \( F(x_n) \neq y \). With the help of the identity (2.3) and the fact \( J_p(z_{n,k}) = J_p(x_n) + u_{n,k} \), we have
\[
\Delta_p(x^i, z_{n,k+1}) - \Delta_p(x^i, z_{n,k}) = \Delta_p(z_{n,k}, z_{n,k+1}) + \langle u_{n,k+1} - u_{n,k}, z_{n,k} - x^i \rangle.
\]
(3.3)

We estimate the two terms on the right-hand side separately. By using the definition of \( u_{n,k+1} \), we have
\[
\langle u_{n,k+1} - u_{n,k}, z_{n,k} - x^i \rangle = \omega_{n,k} \left( f_r(y - F(x_n) - T_n(z_{n,k} - x_n)), T_n(z_{n,k} - x^i) \right).
\]

Writing
\[
T_n(z_{n,k} - x^i) = [y - F(x_n) + T_n(x_n - x^i)] - [y - F(x_n) - T_n(z_{n,k} - x_n)]
\]
and using the properties of the mapping \( f_r \), we therefore obtain
\[
\langle u_{n,k+1} - u_{n,k}, z_{n,k} - x^i \rangle = -\omega_{n,k} \left( f_r(y - F(x_n) - T_n(z_{n,k} - x_n)), y - F(x_n) - T_n(z_{n,k} - x_n) \right)
\]
\[
+ \omega_{n,k} \left( f_r(y - F(x_n) - T_n(z_{n,k} - x_n)), y - F(x_n) + T_n(x_n - x^i) \right)
\]
\[
\leq -\omega_{n,k} \| y - F(x_n) - T_n(z_{n,k} - x_n) \|^r
\]
\[
+ \omega_{n,k} \| y - F(x_n) - T_n(z_{n,k} - x_n) \|^{r-1} \| y - F(x_n) + T_n(x_n - x^i) \|.
\]

By using assumption 3.1, we can obtain
\[
\langle u_{n,k+1} - u_{n,k}, z_{n,k} - x^i \rangle \leq \eta\omega_{n,k} \| y - F(x_n) - T_n(z_{n,k} - x_n) \|^{r-1} \| y - F(x_n) \| - \omega_{n,k} \| y - F(x_n) - T_n(z_{n,k} - x_n) \|^{r}.
\]

According to the definition of \( k_n \) we have
\[
\| y - F(x_n) - T_n(z_{n,k} - x_n) \| > \mu \| y - F(x_n) \|, \quad k = 0, \ldots, k_n - 1.
\]
(3.4)

Consequently,
\[
\langle u_{n,k+1} - u_{n,k}, z_{n,k} - x^i \rangle \leq - \left( 1 - \frac{\eta}{\mu} \right) \omega_{n,k} \| y - F(x_n) - T_n(z_{n,k} - x_n) \|^r.
\]
(3.5)

Next we estimate \( \Delta_p(z_{n,k}, z_{n,k+1}) \). By using (2.4) and \( 1/p + 1/p^* = 1 \), we have
\[
\Delta_p(z_{n,k}, z_{n,k+1}) = \frac{1}{p^*} \left( \| z_{n,k+1} \|^p - \| z_{n,k} \|^p + \| z_{n,k} \|^p - \langle J_p(z_{n,k+1}), z_{n,k} \rangle \right).
\]

Hence, with the help of the properties of \( J_p \) and the fact \( J_p(z_{n,k}) = J_p(x_n) + u_{n,k} \), we obtain
\[
\Delta_p(z_{n,k}, z_{n,k+1}) = \frac{1}{p^*} \left( \| J_p(z_{n,k}) \|^{p^*} - \| J_p(z_{n,k+1}) \|^{p^*} - \langle J_p(z_{n,k+1}) - J_p(z_{n,k}), z_{n,k} \rangle \right)
\]
\[
= \frac{1}{p^*} \left( \| J_p(x_n) \|^{p^*} + u_{n,k+1} \|^p - \| J_p(x_n) + u_{n,k} \|^p \right)
\]
\[
- \langle u_{n,k+1} - u_{n,k}, J_p(x_n) + u_{n,k} \rangle).
\]

Since \( X \) is \( s \)-convex, the dual space \( X^* \) is \( s^* \)-smooth. Thus we can apply the inequality (2.2) to derive that
\[
\Delta_p(z_{n,k}, z_{n,k+1}) \leq C_{p,s^*} \| u_{n,k+1} - u_{n,k} \|^s \left( \| J_p(z_{n,k}) \|^{p^*-s^*} + \| u_{n,k+1} - u_{n,k} \|^{p^*-s} \right)
\]
\[
= C_{p,s^*} \left( \| z_{n,k} \|^{(p-1)(p^*-s)} \| u_{n,k+1} - u_{n,k} \|^{s} + \| u_{n,k+1} - u_{n,k} \|^{p^*-s^*} \right).
\]
This together with the definition of \( u_{n,k+1} \) and the property of \( f_r \) gives
\[
\Delta_p(z_{n,k}, z_{n,k+1}) \leq C_{r,s} \omega_{n,k}^r \|T_n\|^r \|z_{n,k}\|^{(p-1)(p'-s)} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^{(r-1)p'} + C_{r,s} \omega_{n,k}^{r'} \|T_n\|^{r'} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^{(r-1)p'}. \tag{3.6}
\]
Combining the estimates (3.5) and (3.6) with (3.3), we obtain
\[
\Delta_p(x^\dagger, z_{n,k+1}) - \Delta_p(x^\dagger, z_{n,k}) \\
\leq -(1 - \frac{\eta}{\mu}) \omega_{n,k} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^r \\
+ C_{r,s} \omega_{n,k}^r \|T_n\|^r \|z_{n,k}\|^{(p-1)(p'-s)} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^{(r-1)p'} \\
+ C_{r,s} \omega_{n,k}^{r'} \|T_n\|^{r'} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^{(r-1)p'}.
\]
According to the choice of \( \omega_{n,k} \), we can obtain
\[
\Delta_p(x^\dagger, z_{n,k+1}) - \Delta_p(x^\dagger, z_{n,k}) \leq -c_0 \omega_{n,k} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^r,
\]
where \( c_0 > 0 \) is the constant defined by (3.1). This implies (3.2) and
\[
c_0 \sum_{k=0}^l \omega_{n,k} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^r \leq \Delta_p(x^\dagger, x_n) - \Delta_p(x^\dagger, z_{n,l+1}) \tag{3.7}
\]
for all \( l < k_n \). In order to show that \( k_n \) is finite, we observe that the established monotonicity result implies
\[
\Delta_p(x^\dagger, z_{n,k}) \leq \Delta_p(x^\dagger, z_{n,0}) = \Delta_p(x^\dagger, x_0) \leq \Delta_p(x^\dagger, x_0)
\]
which shows that \( \|z_{n,k}\| \leq B \) for some positive constant \( B \) independent of \( n \) and \( k \). This together with \( p \leq s \), (3.4) and \( y \neq F(x_n) \) implies that
\[
\omega_{n,k} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^r \geq d_n, \tag{3.8}
\]
where
\[
d_n := \min \left\{ \theta_1 \mu^p \|T_n\|^{-p} \|y - F(x_n)\|^p, \theta_2 \mu^p B^{p-s} \|T_n\|^{-s} \|y - F(x_n)\|^s \right\} > 0.
\]
Therefore, it follows from (3.7) that \( c_0 d_n (l + 1) \leq \Delta_p(x^\dagger, x_n) < \infty \) for all \( l < k_n \). This shows that \( k_n \) must be finite and hence \( x_{n+1} \) is well defined. By taking \( l = k_n - 1 \) in (3.7), we obtain
\[
c_0 \sum_{k=0}^{k_n-1} \omega_{n,k} \|y - F(x_n) - T_n(z_{n,k} - x_n)\|^r \leq \Delta_p(x^\dagger, x_n) - \Delta_p(x^\dagger, x_{n+1}). \tag{3.9}
\]
Next we will show that \( \{x_n\} \) is a Cauchy sequence. We need to estimate \( \Delta_p(x_l, x_m) \) for any integers \( 0 \leq m < l < \infty \). From the identity (2.3), it follows that
\[
\Delta_p(x_l, x_m) = \Delta_p(x^\dagger, x_m) - \Delta_p(x^\dagger, x_l) + (J_p(x_l) - J_p(x_m), x_l - x^\dagger). \tag{3.10}
\]
By using the fact \( J_p(x_{n+1}) = J_p(x_n) + u_{n,k_n} \), we can write
\[
(J_p(x_l) - J_p(x_m), x_l - x^\dagger) = \sum_{l=0}^{l-1} (J_p(x_{n+1}) - J_p(x_n), x_l - x^\dagger) \\
= \sum_{l=0}^{l-1} (u_{n,k_n}, x_l - x^\dagger). \tag{3.11}
\]
Since \( u_{n,0} = 0 \), we have from the definition of \( u_{n,k} \) and the property of \( j_r \) that
\[
| \{u_{n,k}, x_t - x^\dagger\} | = \left| \sum_{k=0}^{t-1} (u_{n,k+1} - u_{n,k}, x_t - x^\dagger) \right|
= \sum_{k=0}^{t-1} \omega_{n,k} \langle j_r(y - F(x_n) - T_n(z_{a,k} - x_n)), T_n(x_t - x^\dagger) \rangle
\leq \sum_{k=0}^{t-1} \omega_{n,k} \| y - F(x_n) - T_n(z_{a,k} - x_n) \|^r \| T_n(x_t - x^\dagger) \|.
\] (3.12)

To estimate \( \| T_n(x_t - x^\dagger) \| \), we write \( \| T_n(x_t - x^\dagger) \| \leq \| T_n(x_n - x^\dagger) \| + \| T_n(x_n - x_t) \| \). By using assumption 3.1, we then obtain
\[
\| T_n(x_t - x^\dagger) \| \leq (1 + \eta) (\| y - F(x_n) \| + \| F(x_t) - F(x_n) \|)
\leq (1 + \eta) (2\| y - F(x_n) \| + \| y - F(x_t) \|).
\]

By the definition of \( k_n \), we have
\[
\| y - F(x_n) - T_n(x_{n+1} - x_n) \| \leq \mu \| y - F(x_n) \|.
\]

This together with assumption 3.1 implies that
\[
\| y - F(x_{n+1}) \| \leq \| F(x_{n+1}) - F(x_n) - T_n(x_{n+1} - x_n) \|
+ \| y - F(x_n) - T_n(x_{n+1} - x_n) \|
\leq \eta \| F(x_{n+1}) - F(x_n) \| + \mu \| y - F(x_n) \|.
\]

Therefore
\[
\| y - F(x_{n+1}) \| \leq \frac{\mu + \eta}{1 - \eta} \| y - F(x_n) \|.
\]

Since \( (\mu + \eta)/(1 - \eta) < 1 \), this shows that \( \| y - F(x_n) \| \) monotonically decreases to 0.

Consequently, we have for \( n < l \) and \( k = 0, \ldots, k_n - 1 \) that
\[
\| T_n(x_t - x^\dagger) \| \leq 3(1 + \eta) \| y - F(x_n) \| \leq \frac{3(1 + \eta)}{\mu} \| y - F(x_n) - T_n(z_{a,k} - x_n) \|.
\]

Combining this with (3.12) and (3.9) gives
\[
| \{u_{n,k}, x_t - x^\dagger\} | \leq \frac{3(1 + \eta)}{\mu} \sum_{k=0}^{t-1} \omega_{n,k} \| y - F(x_n) - T_n(z_{a,k} - x_n) \|^r
\leq c_1 (\Delta_p(x^\dagger, x_n) - \Delta_p(x^\dagger, x_{n+1})),
\]
where \( c_1 := 3(1 + \eta)/(\mu c_0) \). This together with (3.11) gives
\[
| (J_p(x_t) - J_p(x_n), x_t - x^\dagger) | \leq c_1 (\Delta_p(x^\dagger, x_m) - \Delta_p(x^\dagger, x_t)).
\]

Plugging this inequality into (3.10) yields
\[
\Delta_p(x_t, x_m) \leq (1 + c_1) (\Delta_p(x^\dagger, x_m) - \Delta_p(x^\dagger, x_t)).
\]

Since \( \{\Delta_p(x^\dagger, x_m)\} \) is monotonically decreasing, it is convergent. Thus \( \Delta_p(x_t, x_m) \to 0 \) as \( m, l \to \infty \).

Since \( X \) is uniformly convex, it follows from [16, theorem 2.12(e)] that \( \{x_n\} \) is a Cauchy sequence and thus \( x_m \to x^* \) as \( n \to \infty \) for some \( x^* \in B_{\beta}(x^\dagger, \Delta_p) \subset X \).

Since \( \| F(x_n) - y \| \to 0 \) as \( n \to \infty \) and \( F \) is continuous, we have \( F(x^*) = y \) which implies that \( x^* \) is a solution and thus \( \{x_n\} \) converges to a solution of (1.1). \( \square \)
3.2. Approximate data case

We next consider the case that both the data \( y \) and the operator \( F \) are given approximately. We assume that \( y^\delta \) is the available noisy data for \( y \) satisfying

\[
\| y^\delta - y \| \leq \delta
\]

with a given small noise level \( \delta > 0 \). We also assume that \( F_h : \mathcal{D}(F) \subset \mathcal{X} \to \mathcal{Y} \) and \( G_h : \mathcal{D}(F) \subset \mathcal{X} \to L(\mathcal{X}, \mathcal{Y}) \) are available such that

\[
\| F_h(x) - F(x) \| \leq \xi_h \quad \text{and} \quad \| G_h(x) - F'(x) \| \leq \gamma_h
\]

for all \( x \in B_p(x^\dagger, \Delta_p) \), where \( \xi_h, \gamma_h \to 0 \) as \( h \to 0 \). In this situation, the inexact Newton–Landweber iteration method in Banach spaces will be modified as follows. We first choose an initial guess \( x_0 \in \mathcal{D}(F) \subset \mathcal{X} \). Assume that \( x_n \) is defined. We construct a sequence \( \{u_{n,k}\} \subset \mathcal{X}^* \) iteratively by setting \( u_{n,0} = 0 \) and

\[
u_{n,k+1} = u_{n,k} + \omega_{n,k}G_h(x_n)^*n^j_y(y^\delta - F_h(x_n) - G_h(x_n)(z_{n,k} - x_n)),\]

where \( \omega_{n,k} \) is a positive number and \( z_{n,k} = J_{p^*}(J_p(x_n) + u_{n,k}) \). Let \( k_n \) be the first integer such that

\[
\| y^\delta - F_h(x_n) - G_h(x_n)(z_{n,k_n} - x_n) \| \leq \mu \| y^\delta - F_h(x_n) \|
\]

where \( 0 < \mu < 1 \) is a preassigned number. We then define the next iterate by \( x_{n+1} = z_{n,k_n} \). The outer Newton iteration is then terminated by the discrepancy principle

\[
\| y^\delta - F_h(x_{n(\delta,h)}) \| \leq \tau (\delta + \xi_h + R\gamma_h) < \| y^\delta - F_h(x_n) \|, \quad 0 \leq n < n(\delta, h)
\]

for some given number \( \tau > 1 \), where \( R > 0 \) is a number such that

\[
B_p(x^\dagger, \Delta_p) \subset \{ x \in \mathcal{X} : \| x - x^\dagger \| \leq R \}
\]

which is always possible. This outputs a stopping index \( n(\delta, h) \) and hence an approximate solution \( x_{n(\delta,h)} \). We will consider the approximation property of \( x_{n(\delta,h)} \) to a solution of (1.1) as \( \delta, h \to 0 \).

**Theorem 3.2.** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be Banach spaces with \( \mathcal{X} \) being uniformly smooth and \( s \)-convex for some \( s > 1 \); let assumption 3.1 hold with \( 0 \leq \eta < 1 \), \( x_0 \in B_p(x^\dagger, \Delta_p) \), and \( 1 < p \leq s \) and \( r \geq 1 \) in the definition of \( \{x_n\} \). Then the method is well defined if \( \eta < \mu < 1, \tau > (1+\eta)/\mu(\eta-\eta) \) and \( \omega_{n,k} \) is chosen as

\[
\omega_{n,k} = \min \left\{ \omega_{n,k}^{(1)}, \omega_{n,k}^{(2)} \right\}
\]

where

\[
\omega_{n,k}^{(1)} = \theta_1 \| G_h(x_n) \|^{-p} \| y^\delta - F_h(x_n) - G_h(x_n)(z_{n,k} - x_n) \|^{-p-r},
\]

\[
\omega_{n,k}^{(2)} = \theta_2 (G_h(x_n))^{-r} \| z_{n,k} \|^{-p-r} \| y^\delta - F_h(x_n) - G_h(x_n)(z_{n,k} - x_n) \|^{-r-r}
\]

for some positive constants \( \theta_1 \) and \( \theta_2 \) satisfying

\[
c_2 := 1 - \frac{\eta}{\mu} - \frac{1+\eta}{\mu\tau} - C_{p,r,s} \left( \theta_1^{r-1} + \theta_2^{r-1} \right) > 0.
\]

*If in addition \( F \) is weakly closed, then for any sequences \( \{y^\delta\}, \{F_h\} \) and \( \{G_h\} \) with \( \delta, h \to 0 \) as \( j \to \infty \), the sequence \( \{x_{n(\delta,h)}\} \) contains a subsequence that converges weakly to a solution of (1.1) in \( B_p(x^\dagger, \Delta_p) \). If \( x^\dagger \) is the unique solution of (1.1) in \( B_p(x^\dagger, \Delta_p) \), then \( x_{n(\delta,h)} \) converges weakly to \( x^\dagger \) as \( \delta, h \to 0 \).*
\textbf{Proof.} We use the similar argument in the proof of theorem 3.1 with suitable modifications and show that each \( k_n \) is finite, \([\Delta_p(x^1, x_n)]\) is monotonically decreasing and \( n(\delta, h) < \infty \).

Similarly to the derivation of (3.6), we can obtain
\[
\Delta_p(z_{n,k}, z_{n,k+1}) \\
\leq C_{p,r} \omega_{n,k} \|G_h(x_n)\| p \|z_{n,k}\| (p-1)(p-r') \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| (r-1)p' \\
+ C_{p,r} \omega_{n,k} \|G_h(x_n)\| p \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| (r-1)p'.
\]

Moreover, we can also derive
\[
\{u_{n,k+1} - u_{n,k}, z_{n,k} - x^1\} \\
\leq \omega_{n,k} \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| (r-1) \|y^\delta - F_h(x_n)\| \\
- \omega_{n,k} \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| r'.
\]

By using (3.13), (3.14), assumption 3.1 and the fact
\[
\|y^\delta - F_h(x_n)\| > \tau (\delta + \xi_k + R\gamma h), \quad 0 < n < n(\delta, h),
\]
we have
\[
\|y^\delta - F_h(x_n)\| < \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\|, \quad k = 0, \ldots, k_n - 1
\]
gives
\[
\{u_{n,k+1} - u_{n,k}, z_{n,k} - x^1\} \\
\leq - \left(1 - \frac{1}{\tau \mu} (1 + \eta + \tau \eta) \right) \omega_{n,k} \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| r'.
\]

Therefore, it follows from identity (2.3) that
\[
\Delta_p(x^1, z_{n,k+1}) - \Delta_p(x^1, z_{n,k}) \\
\leq - \left(1 - \frac{1}{\tau \mu} (1 + \eta + \tau \eta) \right) \omega_{n,k} \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| r' \\
+ C_{p,r} \omega_{n,k} \|G_h(x_n)\| p \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| (r-1)p' \\
+ C_{p,r} \omega_{n,k} \|G_h(x_n)\| p \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| (r-1)p'.
\]

According to the choice of \( \omega_{n,k} \), we obtain
\[
\Delta_p(x^1, z_{n,k+1}) - \Delta_p(x^1, z_{n,k}) \leq -c_2 \omega_{n,k} \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| r' \quad (3.19)
\]
for \( k = 0, \ldots, k_n - 1 \), where \( c_2 > 0 \) is the constant defined by (3.16). This implies that
\[
\Delta_p(x^1, z_{n,k+1}) \leq \Delta_p(x^1, z_{n,k}).
\]

Moreover, similar to the derivation of (3.8) we can obtain
\[
\omega_{n,k} \|y^\delta - F_h(x_n) - G_h(x_n) (z_{n,k} - x_n)\| \geq d_\delta(\delta, h),
\]
where
\[
d_\delta(\delta, h) := C \min \{\|G_h(x_n)\|^{-p} \|y^\delta - F_h(x_n)\|, \|G_h(x_n)\|^{-1} \|y^\delta - F_h(x_n)\| \}
\]
for some positive constant \( C \). Thus, it follows from (3.19) that
\[
c_2 (l + 1) d_\delta(\delta, h) \leq \Delta_p(x^1, x_n) - \Delta_p(x^1, x_{n+1}), \quad 0 \leq l < k_n.
\]
By using (3.17) we can see \( d_n(\delta, h) > 0 \). Therefore \( k_n \) must be finite and
\[
c_{2}k_{n}d_{n}(\delta, h) \leq \Delta_{p}(x^{n}, x_{n}) - \Delta_{p}(x^{1}, x_{n+1}).
\]
Summing over \( n \) from 0 to \( n(\delta, h) - 1 \), we obtain
\[
c_{2} \sum_{n=0}^{n(\delta, h)-1} k_{n}d_{n}(\delta, h) \leq \Delta_{p}(x_{0}, x^{1}) < \infty.
\]
(3.20)
Observe that assumption 3.1 and (3.14) imply that \( \|G_{h}(x_{n})\| \) can be bounded by a constant independent of \( n \). Since (3.17) holds, \( d_{n}(\delta, h) \) can be bounded below by a positive number independent of \( n \). Since \( k_{n} \geq 1 \), we can conclude from (3.20) that \( n(\delta, h) < \infty \). The method is thus well defined. Moreover, \( \Delta_{p}(x^{1}, x_{n(\delta, h)}) \leq \Delta_{p}(x^{1}, x_{0}) \) and thus \( \{x_{n(\delta, h)}\} \) is bounded in \( \mathcal{X} \). Recall that
\[
\|y^{\delta} - F_{h}(x_{n(\delta, h)})\| \leq \tau(\delta + \xi_{h} + R\gamma_{h}),
\]
and by using (3.13) and (3.14) we have \( F(x_{n(\delta, h)}) \rightarrow y \) as \( \delta, h \rightarrow 0 \). Since \( \mathcal{X} \) is reflexive and \( F \) is weakly closed, the remaining part can be done by a standard argument. \( \square \)

**Remark 3.3.** We point out that the method in fact terminates after \( O(1 + |\log(\delta + \xi_{h} + \gamma_{h})|) \) outer iterations if \( 0 < \eta < 1/3, \eta < \mu < 1 - 2\eta \) and \( \tau > 2(1 + \eta)/(1 - 2\eta - \mu) \). We can argue this as in [3]. Recall that
\[
\|y^{\delta} - F_{h}(x_{n}) - G_{h}(x_{n})(x_{n+1} - x_{n})\| \leq \mu\|y^{\delta} - F_{h}(x_{n})\|.
\]
From (3.14) and assumption 3.1, it then follows for \( 0 \leq n < n(\delta, h) \) that
\[
\|y^{\delta} - F_{h}(x_{n+1})\| \leq \mu\|y^{\delta} - F_{h}(x_{n})\| + \|F_{h}(x_{n+1}) - F_{h}(x_{n}) - G_{h}(x_{n})(x_{n+1} - x_{n})\|
\leq \mu\|y^{\delta} - F_{h}(x_{n})\| + 2(\xi_{h} + R\gamma_{h}) + \eta\|F_{h}(x_{n+1}) - F_{h}(x_{n})\|
\leq \mu\|y^{\delta} - F_{h}(x_{n})\| + 2(1 + \eta)(\xi_{h} + R\gamma_{h}) + \eta\|F_{h}(x_{n+1}) - F_{h}(x_{n})\|.
\]
This together with (3.17) implies that
\[
\|y^{\delta} - F_{h}(x_{n+1})\| \leq \beta\|y^{\delta} - F_{h}(x_{n})\|, \quad n = 0, \ldots, n(\delta, h) - 1,
\]
where \( \beta := [\mu + \eta + 2(1 + \eta)/\tau]/(1 - \eta) < 1 \). Therefore
\[
\tau(\delta + \xi_{h} + R\gamma_{h}) < \|y^{\delta} - F_{h}(x_{n(\delta, h) - 1})\| \leq \beta^{n(\delta, h) - 1}\|y^{\delta} - F_{h}(x_{0})\|
\]
which shows that \( n(\delta, h) \leq O(1 + |\log(\delta + \xi_{h} + \gamma_{h})|) \).

**Remark 3.4.** In theorem 3.2, we only obtain the weak convergence. The proof of strong convergence remains open even in the Hilbert space setting. However, our numerical results strongly suggest that the strong convergence holds. On the other hand, in some situations we are interested in the strong convergence in a Banach space \( Z \) in which \( \mathcal{X} \) can be compactly embedded; the weak convergence in \( \mathcal{X} \) is already enough for the purpose. For instance, consider the identification of the parameter \( a \) in
\[
\begin{align*}
\text{div}(a \nabla u) &= f & \text{in } \Omega, \\
u &= g & \text{on } \partial \Omega,
\end{align*}
\]
(3.21)
from an \( L^{2} \) measurement of \( u \), where \( \Omega \subset \mathbb{R}^{d}, d = 2, 3 \), is a bounded domain with smooth boundary \( \partial \Omega, f \in H^{-1}(\Omega) \) and \( g \in H^{1/2}(\partial \Omega) \). It is well known that for \( a \in L^{\infty}(\Omega) \) with \( a \geq \nu_{0} > 0 \) on \( \Omega \), (3.21) has a unique solution \( u = u(a) \in H^{1}(\Omega) \). Thus the inverse problem requires inverting the map \( F : L^{\infty}(\Omega) \rightarrow L^{2}(\Omega) \) with \( F(a) := u(a) \). Since \( L^{\infty}(\Omega) \) is not convenient to use, we may consider \( F \) on a smaller space. When formulating the problem in the framework of Hilbert spaces, we may consider \( F \) over \( H^{2}(\Omega) \) which, however, makes
the calculation of the Fréchet derivatives and their adjoints rather complicated. It seems more natural to consider $F$ on the Sobolev space $W^{1,p}(\Omega)$ with $p > d$. Assumption 3.1 has been verified in [3] for $F$ being an operator from $W^{1,p}(\Omega)$ to $L^2(\Omega)$. Thus, theorem 3.2 implies the weak convergence of our method in $W^{1,p}(\Omega)$ which in turn implies the strong convergence in $L^\infty(\Omega)$ since $W^{1,p}(\Omega)$ embeds compactly into $L^\infty(\Omega)$ for $p > d$.

### 4. Numerical experiments

In this section, we present some numerical results to test the inexact Newton–Landweber method in Banach spaces. We consider the estimation of the coefficient $c$ in the boundary value problem

$$
\begin{cases}
-u'' + cu = f & \text{in } (0, 1) \\
 u(0) = g_0, & u(1) = g_1
\end{cases}
$$

(4.1)

from the measurement of the state variable $u$, where $g_0, g_1$ and $f \in H^{-1}[0, 1]$ are given. It is well known that (4.1) has a unique solution $u := u(c) \in H^1[0, 1]$ for each $c$ in the domain

$$
\mathcal{D} := \{c \in L^\infty[0, 1] : \|c - \hat{c}\|_{L^p} \leq \gamma_p \text{ for some } \hat{c} \geq 0 \text{ a.e.}\}
$$

with some $\gamma_p > 0$, where $1 \leq p \leq \infty$. For our numerical simulations, we take $\mathcal{X} = L^p[0, 1]$ and $\mathcal{Y} = L^q[0, 1]$ with $1 < p < \infty$ and $1 \leq r \leq \infty$, and identify $c \in \mathcal{X}$ from an $L^p[0, 1]$-measurement $u'$ of $u$. Thus the inverse problem reduces to solving (1.1) with the nonlinear operator $F : \mathcal{D} \subset L^p[0, 1] \mapsto L^q[0, 1]$ defined as $F(c) := u(c)$. It is easy to show that $F$ is Fréchet differentiable, and the Fréchet derivative and its Banach space adjoint are given by

$$
F'(c)h = -A(c)^{-1}(hu(c)), \quad F'(c)^*w = -u(c)A(c)^{-1}w,
$$

where $A(c) : H^2 \cap H^1_0 \mapsto L^2$ is defined by $A(c)u = -u'' + cu$. Recall that in the space $L^p[0, 1]$ with $1 < p < \infty$, the duality mapping $J_p : L^p[0, 1] \mapsto L^{p'}[0, 1]$ is given by

$$
J_p(c) = |c|^{p-1}\text{sign}(c), \quad c \in L^p[0, 1].
$$

In our numerical computation, all differential equations are solved approximately by the finite difference method by dividing the interval $[0, 1]$ into $N + 1$ subintervals with equal length $h = 1/(N + 1)$; we take $N = 400$ in our actual computation.

**Example 4.1.** We first test our method for sparsity reconstruction. We assume that the parameter $c$ in (4.1) to be reconstructed is

$$
c^\dagger(t) = \begin{cases} 
0.5, & 0.3 \leq t \leq 0.4, \\
1.0, & 0.6 \leq t \leq 0.7, \\
0, & \text{elsewhere}
\end{cases}
$$

by assuming that $u(c^\dagger) = 1 + 5t, f(t) = (1 + 5t)c^\dagger(t), g_0 = 1$ and $g_1 = 6$. In our computation, instead of $u(c^\dagger)$ we use random noise data $u'$ satisfying $\|u' - u(c^\dagger)\|_{L^2[0,1]} = \delta$ with noise level $\delta > 0$. When applying the method proposed in section 3, we take $\mu = 0.96$, $\tau = 1.05$, $\theta_1 = 0.40$ and $\theta_2 = 0.40$; we also take the initial guess to be $c_0 = 0$. Figure 1(a) reports the result via the inexact Newton–Landweber method in Hilbert spaces with $\mathcal{X} = L^2[0, 1]$ and $\mathcal{Y} = L^2[0, 1]$. It is clear that the reconstructed solution is oscillatory at the zero parts which destroys the sparsity of the exact solution $c^\dagger$. In figure 1(b), we report the computational result by our method with $\mathcal{Y} = L^2[0, 1]$ and $\mathcal{X} = L^p[0, 1]$ for $p > 1$ but close to 1; we take $p = 1.1$. The reconstruction of the sparsity is significantly improved. However, we have to pay the price that the reconstruction is oscillatory on the non-zero part which is typical for $L^p$-regularization with $p \geq 1$ close to 1. In order to remove this notorious effect, a possible
way is to develop a method with the duality mapping $J_p$ replaced by suitable mappings related to the total variational like functionals.

**Example 4.2.** We next give a test result on our method when the noisy data contain some data points that are highly inconsistent with other data points. Such data points are called outliers, and may arise from procedural measurement error.

We estimate again the parameter $c$ in (4.1) by assuming that $f(t) = (1 - 2t)(4 - t + 3 \sin(2\pi t))$, $g_0 = 1$ and $g_1 = -1$. If $u(c^\dagger) = 1 - 2t$, then

$$c^\dagger(t) = 4 - t + 3 \sin(2\pi t)$$

is the sought solution. In figure 2, we present the reconstruction results by the method in section 3 with $\tau = 1.1$, $\mu = 0.92$, $\theta_1 = \theta_2 = 0.4$ and the initial guess $c_0 = 4 - t$. Figures 2(a) and (d) present the plots of the noisy data; the one in (a) contains only Gaussian noise, while the one in (d) contains not only Gaussian noise but also a few outliers. Figures 2(b) and (e) present the reconstruction results by the inexact Newton–Landweber method with $X = Y = L^2[0, 1]$. It is clear that the method is highly susceptible to even small number of outliers. In figures 2(c) and (f) we present reconstruction results by our method with $X = L^2[0, 1]$ and $Y = L^2[0, 1]$ with $r = 1.1$. It can be seen that the method is robust enough to prevent being affected by outliers.

We also performed some numerical tests on our method with different choices of the initial guess and found that the convergence speed is affected significantly. This suggests that one may consider the rate of convergence by formulating a suitable source condition on $c_0 - c^\dagger$. 

---

**Figure 1.** Numerical results for example 4.1 with noise level $\delta = 0.1 \times 10^{-3}$: (a) $X = Y = L^2[0, 1]$, (b) $X = L^1[0, 1]$ and $Y = L^2[0, 1]$. 

---

![Figure 1](image-url)
The order optimality of the inexact Newton–Landweber method in Hilbert spaces has been established in [7] recently. The corresponding question in Banach spaces, however, is rather challenging.

Acknowledgment

This work is partially supported by the DECRA grant DE120101707 of Australian Research Council.

References


