Nonstationary iterated Tikhonov regularization for ill-posed problems in Banach spaces

This article has been downloaded from IOPscience. Please scroll down to see the full text article.
2012 Inverse Problems 28 104011
(http://iopscience.iop.org/0266-5611/28/10/104011)

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

Download details:
IP Address: 203.129.48.138
The article was downloaded on 03/10/2012 at 02:02

Please note that terms and conditions apply.
Nonstationary iterated Tikhonov regularization for ill-posed problems in Banach spaces

Qinian Jin and Linda Stals

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

Received 16 November 2011, in final form 15 February 2012
Published 1 October 2012
Online at stacks.iop.org/IP/28/104011

Abstract

Nonstationary iterated Tikhonov regularization is an efficient method for solving ill-posed problems in Hilbert spaces. However, this method may not produce good results in some situations since it tends to oversmooth solutions and hence destroy special features such as sparsity and discontinuity. By making use of duality mappings and Bregman distance, we propose an extension of this method to the Banach space setting and establish its convergence. We also present numerical simulations which indicate that the method in Banach space setting can produce better results.

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

1. Introduction

We consider the linear operator equation

$$Ax = y,$$

(1.1)

where $A : \mathcal{X} \rightarrow \mathcal{Y}$ is a bounded linear operator between two Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ with a non-closed range. We will assume (1.1) has a solution. Such equations are, in general, ill-posed in the sense that their solutions do not depend continuously on the data. Let $y^\delta$ be noisy data satisfying

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

(1.2)

with a given small noise level $\delta > 0$. We will use $y^\delta$ to construct a stable approximate solution to (1.1).

Many continuous and iterative regularization methods have been developed to solve (1.1) stably when both $\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces; see [5]. The most well-known method is Tikhonov regularization which defines the regularized solutions by

$$x^\alpha_u = (\alpha I + A^*A)^{-1}A^*y^\delta,$$

where $\alpha > 0$ is a regularization parameter.
where $\alpha > 0$ is the regularization parameter whose choice strongly influences the approximation property of the method, see [8]. For practical applications, the regularization parameter $\alpha$ is usually determined by a posteriori rules. Thus, the implementation of Tikhonov regularization requires the computing of the regularized solutions for several different values of $\alpha$ in order to update a good approximate solution. During the process, the computed regularized solutions are not fully used in subsequent computation. The nonstationary iterated Tikhonov regularization is an attractive variant which defines the regularized solutions iteratively by

\[ x_\delta^n = x_{n-1}^\delta - (\alpha_n I + A^* A)^{-1} A^* (Ax_{n-1}^\delta - y^\delta), \]

where $x_0^\delta := x_0 \in \mathcal{X}$ is an initial guess and $\{\alpha_n\}$ is a preassigned sequence of positive numbers.

It has been shown that, when the sequence $\{\alpha_n\}$ satisfies suitable property and the discrepancy principle is used to terminate the iteration, the regularization property of this method has been established; see [9] and references therein. It is worthwhile to point out that each $x_\delta^n$ is the unique minimizer of the functional

\[ x \rightarrow \frac{1}{2} \|Ax - y^\delta\|^2 + \frac{\alpha_n}{2} \|x - x_{n-1}^\delta\|^2 \text{ over } \mathcal{X}. \]

This method has also been extended to solve nonlinear inverse problems in Hilbert spaces [13].

Regularization methods in Hilbert spaces can produce a good result when the sought solution is smooth and the data contain only Gaussian noise. Since such methods have a tendency to oversmooth solutions and are susceptible to the types of noise, they may not produce good results in some applications. Instead, regularization in Banach spaces can be used to obtain better reconstructions. For instance, in sparsity recovery the $L^1$-penalty, or sometimes the $L^p$-penalty with $p > 1$ sufficiently close to 1, is used in variational regularization methods [18]; the $L^p$-penalty with $0 < p < 1$ is even used in this situation [7]. In imaging analysis, the total variational (TV) penalty term was introduced in [16] to produce striking results. In some statistical situations, due to procedural error the measurement data may contain data points that are highly inconsistent with other data points, the $L^2$ regression cannot produce a satisfactory estimator; however, the $L^1$ regression can give a rather robust estimator [1, 15].

Moreover, in many applications, it is more natural to formulate the inverse problems in Banach spaces rather than in Hilbert spaces. Therefore, it is important to develop stable algorithms to solve inverse problems in the framework of Banach spaces.

Due to its variational formulation, Tikhonov regularization can be easily adapted to solve inverse problems in Banach spaces and extensive work, including the convergence and the derivation of convergence rates, has been carried out; see [3, 10, 14] and references therein. There are also some attempts to extend iterative regularization methods from Hilbert space setting to Banach space setting. For instance, by using duality mappings, Landweber iteration has been formulated to solve ill-posed inverse problems in Banach spaces, see [17, 11].

In this paper, we will extend the nonstationary iterated Tikhonov regularization to Banach space setting. Considering the variation property (1.3), it seems natural to formulate the method in Banach spaces as

\[ x_\delta^n := \arg \min_{x \in \mathcal{X}} \left\{ \frac{1}{r} \|Ax - y^\delta\|^r + \frac{\alpha_n}{p} \|x - x_{n-1}^\delta\|^p \right\}, \]

where $1 \leq p, r < \infty$ are two numbers. Unfortunately, it is difficult to carry out the convergence analysis for this scheme. Therefore, we will modify this scheme and replace the penalty term $\frac{\alpha_n}{p} \|x - x_{n-1}^\delta\|^p$ by a suitable functional which reduces to $\frac{1}{2} \|x - x_{n-1}^\delta\|^2$ when $\mathcal{X}$ is a Hilbert space and $p = 2$. We will achieve this by making use of duality mappings and the Bregman distance.
This paper is organized as follows. In section 2, we will give some preliminary results on Banach spaces. In section 3, we then formulate the nonstationary iterated Tikhonov regularization in Banach spaces and establish the convergence result. In section 4, we will present numerical results to indicate the advantage of the method. Finally, in section 5 we will conclude this paper with several questions that might be interesting for further study.

2. Preliminaries on Banach spaces

Let $X$ be a Banach space with norm $\| \cdot \|$. We will use $X^\ast$ to denote its dual space. Given $x \in X$ and $x^\ast \in X^\ast$ we will write $\langle x^\ast, x \rangle = x^\ast(x)$ for the duality pair. We will use `$\rightharpoonup$' and `$\to$' to denote the strong convergence and the weak convergence, respectively. If $Y$ is another Banach space and $A : X \to Y$ is a bounded linear operator, we will use $A^\ast : Y^\ast \to X^\ast$ to denote its dual, i.e. $\langle A^\ast y^\ast, x \rangle = \langle y^\ast, Ax \rangle$ for any $x \in X$ and $y^\ast \in Y^\ast$.

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

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

is called the duality mapping with gauge function $t \to t^{p-1}$. $J_p$ is in general multi-valued and equals the subdifferential of the convex functional $x \to \| x \|^p/p$. The mapping $J_p$ is monotone, i.e.

$$\langle x^\ast_1 - x^\ast_2, x_1 - x_2 \rangle \geq 0$$

for $x_1, x_2 \in X$, $x^\ast_1 \in J_p(x_1)$ and $x^\ast_2 \in J_p(x_2)$.

The duality mappings have nice properties when $X$ has favorable geometric features. A Banach space $X$ is called strictly convex if for any $x_1, x_2 \in X$ with $x_1 \neq x_2$ and $\| x_1 \| = \| x_2 \| = 1$ there holds $\| x_1 + x_2 \| < 2$, and it is called uniformly convex if its modulus of convexity

$$\delta_X(\epsilon) := \inf \{ 2 - \| x + x^\ast \| : \| x \| = \| x^\ast \| = 1, \| x - x^\ast \| \geq \epsilon \}$$

satisfies $\delta_X(\epsilon) > 0$ for all $0 < \epsilon \leq 2$. Any uniformly convex Banach space is reflexive and strictly convex. Moreover, there holds the following useful result.

**Lemma 2.1.** Let $X$ be a Banach space.

(a) If $X$ is strictly convex, then every duality mapping $J_p$ of $X$ with $1 < p < \infty$ is strictly monotone, i.e. $\langle x^\ast_1 - x^\ast_2, x_1 - x_2 \rangle > 0$ for all $x_1, x_2 \in X$ with $x_1 \neq x_2$ and $x^\ast_1 \in J_p(x_1)$, $x^\ast_2 \in J_p(x_2)$.

(b) If $X$ is uniformly convex, then for any sequence $\{ x_n \} \subset X$ satisfying $x_n \rightharpoonup x$ and $\| x_n \| \to \| x \|$ as $n \to \infty$ there holds $\| x_n - x \| \to 0$ as $n \to \infty$.

A Banach space $X$ is called smooth if for every $x \neq 0$ there is a unique $x^\ast \in X^\ast$ such that $\| x^\ast \| = 1$ and $\langle x^\ast, x \rangle = \| x \|$, and it is called uniformly smooth if its modulus of smoothness

$$\rho_X(s) := \sup \{ \| x + x^\ast \| + \| x - x^\ast \| - 2 : \| x \| = 1, \| x^\ast \| \leq s \}$$

satisfies $\lim_{s \to 0} \frac{\rho_X(s)}{s} = 0$. Any uniformly smooth Banach space is reflexive and smooth. Moreover, a Banach space $X$ is uniformly smooth (resp. uniformly convex) if and only if its dual $X^\ast$ is uniformly convex (resp. uniformly smooth).

**Lemma 2.2.** Let $X$ be a Banach space and let $1 < p < \infty$.

(a) If $X$ is smooth, then every duality mapping $J_p$ of $X$ is single valued. If in addition $X$ is uniformly convex, then every $J_p$ is norm-to-weak continuous.

(b) If $X$ is uniformly smooth, then every $J_p$ is uniformly continuous on bounded subsets of $X$. 

The proofs of lemmas 2.1 and 2.2 can be found in [4], where one can find many other interesting facts on duality mappings together with examples of uniformly smooth and uniformly convex Banach spaces including the sequence spaces $L^p$, the Lebesgue spaces $L^p$, the Sobolev spaces $W^{k,p}$ and the Besov spaces $B^{s,p}$ with $1 < p < \infty$.

In order to formulate the method in Banach spaces and study the convergence property, it is more convenient to use the Bregman distance instead of the norm. The Bregman distance was introduced in [2] for convex functionals in the context of optimization. When it is more convenient to use the Bregman distance instead of the norm. The Bregman distance can be used to get information with respect to the norm. Thus, associated with $J_p$, we can introduce the Bregman distance

$$\Delta_p(x, x) := \frac{1}{p} \|\tilde{x}\|^p - \frac{1}{p} \|x\|^p - \langle J_p(x), \tilde{x} - x \rangle.$$  \hspace{1cm} (2.1)

It is easy to show for any $x, x_1, x_2 \in X$ that

$$\Delta_p(x, x_1) - \Delta_p(x, x_2) = -\Delta_p(x_1, x_2) + \langle J_p(x_1) - J_p(x_2), x_1 - x \rangle.$$  \hspace{1cm} (2.2)

Let $p'$ be the number conjugate to $p$, i.e. $1/p + 1/p' = 1$. Then, by using the properties of the duality mapping $J_p$, we have

$$\Delta_p(x, x) \geq \frac{1}{p} \|\tilde{x}\|^p + \frac{1}{p'} \|x\|^p - \|x\|^{p-1} \|\tilde{x}\|.$$  \hspace{1cm} (2.3)

which implies that $\Delta_p(x, x) \geq 0$ by Young’s inequality. Thus the Bregman distance is non-negative. If $\{x_n\} \subset X$ is a sequence such that $\{\Delta_p(x_n, x)\}$ is bounded, then it follows easily from (2.3) that $\{x_n\}$ is a bounded sequence in $X$.

The Bregman distance is in general not a metric since it does not satisfy the symmetry and the triangle inequality. However, in a smooth and uniformly convex Banach space, the Bregman distance can be used to get information with respect to the norm.

**Lemma 2.3.** Let $X$ be a smooth and uniformly convex Banach space. Then for any $x \in X$ and sequence $\{x_n\} \subset X$, the following hold:

1. $\lim_{n \to \infty} \|x_n - x\| = 0 \iff \lim_{n \to \infty} \Delta_p(x, x_n) = 0 \iff \lim_{n \to \infty} \Delta_p(x_n, x) = 0$.

2. $\{x_n\}$ is a Cauchy sequence if and only if $\Delta_p(x_n, x_m) \to 0$ as $m, n \to \infty$.

The proof of this result, which is based on the characterization of uniformly convex Banach spaces in [19], can be found in [17].

3. **The method and convergence analysis**

We now return to (1.1), where $A : X \to Y$ is a bounded linear operator between two Banach spaces $X$ and $Y$ whose norms are denoted by $\| \cdot \|$. We will always assume that $X$ is a smooth and uniformly convex Banach space. In general, (1.1) may have many solutions. In order to find the desired one, some selection criterion should be addressed. We pick $x_0 \in X$ which may incorporate some available information on the sought solution and define $x^1$ to be the solution of (1.1) with the property

$$\Delta_p(x_0, x_0) = \min \{ \Delta_p(x, x_0) : Ax = y \}.$$ 

By using the uniform convexity of $X$ and the weak lower semi-continuity of norms in Banach spaces, it is easy to see that such $x^1$ exists and is unique. Observe that, for any $z \in N(A)$ the null space of $A$ and $t \in \mathbb{R}$, $x^1 + tz$ are solutions of (1.1). By the minimality of $x^1$, the function $t \mapsto \Delta_p(x^1 + tz, x_0)$ achieves its minimum at $t = 0$. This implies that $x^1$ verifies the equation

$$\langle J_p(x^1) - J_p(x_0), z \rangle = 0, \quad \forall z \in N(A).$$  \hspace{1cm} (3.1)
It is worth pointing out that, when \( x_0 = 0 \), \( x^1 \) is the solution of (1.1) with smallest norm in \( X \).

In order to produce a stable approximation to \( x^1 \) from available noisy data satisfying (1.2), we propose a version of the nonstationary iterated Tikhonov regularization in Banach spaces. We start with \( x_0 \in X \) and a given sequence \( \{\alpha_n\}_{n=1}^{\infty} \) of positive numbers. We define the iterative sequence \( \{x_n^\delta\} \) successively by letting \( x_0^\delta := x_0 \) and letting \( x_n^\delta \) with \( n \geq 1 \) be the minimizer of the convex minimization problem

\[
\min_{x \in X} \left\{ \frac{1}{r} \|Ax - y^\delta\|^r + \alpha_n \Delta_p(x, x_{n-1}^\delta) \right\} \tag{3.2}
\]

whose existence and uniqueness are guaranteed by the uniform convexity of \( X \). We terminate the iteration by the discrepancy principle

\[
\|Ax_n^\delta - y^\delta\| \leq \tau \delta < \|Ax_n^\delta - y^\delta\|, \quad 0 \leq n \leq n_\delta,
\]

with a given number \( \tau > 1 \). This outputs \( x_n^\delta \) which will be used to approximate \( x^1 \).

The key ingredient in the above method is the resolution of the minimization problem (3.2). When \( X \) and \( Y \) are Hilbert spaces and \( p = r = 2 \), (3.2) reduces to (1.3) and thus \( x_n^\delta \) can be written explicitly. In the general Banach space setting, \( x_n^\delta \) does not have an explicit formula which increases the difficulty of convergence analysis. Moreover, finding efficient methods for solving the convex minimization problem (3.2) is indeed an interesting question.

From the definition of \( \{x_n^\delta\} \) it is easy to see that

\[
\|Ax_n^\delta - y^\delta\| \leq \|Ax_{n-1}^\delta - y^\delta\|, \quad n = 1, 2, \ldots \tag{4.4}
\]

Moreover, as the minimizer of the minimization problem (3.2), \( x_n^\delta \) satisfies the optimality condition

\[
\alpha_n (J_p(x_n^\delta) - J_p(x_{n-1}^\delta)) \in A^* \mathcal{J}_r (y^\delta - Ax_n^\delta), \tag{3.5}
\]

where \( \mathcal{J}_r : Y \to 2^Y \) denotes the subdifferential of the convex functional \( z \to \|z\|^r/\r \). When \( 1 < r < \infty \), \( \mathcal{J}_r \) is exactly the duality mapping on \( Y \) with a gauge function \( t \to t^{r-1} \), while for \( r = 1 \), \( \mathcal{J}_1 \) is the indicator function on \( Y^* : \|z\| = 1 \) for \( z \neq 0 \) and \( \mathcal{J}_1(0) = [z^* \in Y^* : \|z^*\| = 1] \). Therefore, for \( 1 \leq r < \infty \), \( z \in Y \) and \( z^* \in \mathcal{J}_r(z) \), there hold

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

where we used the convention \( 0^0 = 1 \).

In the following, we first consider the noise-free iterative sequence \( \{x_n\} \), where each \( x_n \), \( n \geq 1 \), is the minimizer of (3.2) with \( y^\delta \) replaced by \( y \).

**Lemma 3.1.** Let \( X \) and \( Y \) be two Banach spaces with \( X \) being smooth and uniformly convex, \( \{\alpha_n\}_{n=1}^{\infty} \) be a sequence of positive numbers satisfying \( \sum_{n=1}^{\infty} \alpha_n = \infty \) and \( 1 < p < \infty \) and \( 1 \leq r < \infty \) in the definition of \( \{x_n\} \). Then the sequence \( \{x_n\} \) converges to the solution \( x^1 \) of (1.1) as \( n \to \infty \).

**Proof.** We have from (2.2) that

\[
\Delta_p(x^1, x_n) - \Delta_p(x^1, x_{n-1}) = -\Delta_p(x_n, x_{n-1}) + (J_p(x_n) - J_p(x_{n-1}), x_n - x^1).
\]

This together with the non-negativity of the Bregman distance and the optimality condition (3.5) for \( x_n \) gives

\[
\Delta_p(x^1, x_n) - \Delta_p(x^1, x_{n-1}) \leq \frac{1}{\alpha_n} \langle A^* \xi_n, x_n - x^1 \rangle
\]

\[
= \frac{1}{\alpha_n} \langle \xi_n, Ax_n - y \rangle, \tag{3.7}
\]

5
where $\xi_n \in \mathcal{J}_r(y - Ax_n)$. According to property (3.6) of the mapping $\mathcal{J}_r$, we obtain
\[
\Delta_p(x^\dagger, x_n) - \Delta_p(x^\dagger, x_{n-1}) \leq -\frac{1}{\alpha_n} \|Ax_n - y\|'.
\]
This implies for all $n$ that
\[
\Delta_p(x^\dagger, x_n) \leq \Delta_p(x^\dagger, x_{n-1})
\]
and
\[
\frac{1}{\alpha_n} \|Ax_n - y\|' \leq \Delta_p(x^\dagger, x_{n-1}) - \Delta_p(x^\dagger, x_n).
\]
Consequently,
\[
\sum_{j=1}^n \frac{1}{\alpha_j} \|Ax_j - y\|' \leq \Delta_p(x^\dagger, x_0).
\]
By using the monotonicity of $\|Ax_n - y\|$, we obtain
\[
\|Ax_n - y\|' \sum_{j=1}^n \alpha_j^{-1} \leq \Delta_p(x^\dagger, x_0) < \infty.
\]
Since $\sum_{j=1}^n \alpha_j^{-1} \to \infty$ as $n \to \infty$, we must have $\|y - Ax_n\| \to 0$ as $n \to \infty$.

Next we will show that $\{x_n\}$ is a Cauchy sequence. From (2.2) we have for $0 \leq l < m < \infty$ that
\[
\Delta_p(x_m, x_l) = \Delta_p(x^\dagger, x_l) - \Delta_p(x^\dagger, x_m) + (J_p(x_m) - J_p(x_l), x_m - x^\dagger).
\]
We can write
\[
(J_p(x_m) - J_p(x_l), x_m - x^\dagger) = \sum_{n=l+1}^m (J_p(x_n) - J_p(x_{n-1}), x_m - x^\dagger)
\]
\[
= \sum_{n=l+1}^m \frac{1}{\alpha_n} (A^\ast \xi_n, x_m - x^\dagger)
\]
\[
= \sum_{n=l+1}^m \frac{1}{\alpha_n} \langle \xi_n, Ax_m - y \rangle
\]
\[
\leq \sum_{n=l+1}^m \frac{1}{\alpha_n} \|\xi_n\| \|Ax_m - y\|.
\]
Since $\xi_n \in \mathcal{J}_r(y - Ax_n)$, we have from (3.6) that $\|\xi_n\| \leq \|y - Ax_n\|^{r-1}$. Recall also that $\|Ax_m - y\| \leq \|Ax_n - y\|$ for $n \leq m$. We therefore obtain from (3.9) that
\[
(J_p(x_m) - J_p(x_l), x_m - x^\dagger) \leq \sum_{n=l+1}^m \frac{1}{\alpha_n} \|Ax_n - y\|'
\]
\[
\leq \Delta_p(x^\dagger, x_l) - \Delta_p(x^\dagger, x_m).
\]
This together with (3.10) gives
\[
\Delta_p(x_m, x_l) \leq 2(\Delta_p(x^\dagger, x_l) - \Delta_p(x^\dagger, x_m)).
\]
According to the monotonicity (3.8) of $\{\Delta_p(x^\dagger, x_n)\}$, we can conclude that $\Delta_p(x_m, x_l) \to 0$ as $l, m \to \infty$. It then follows from lemma 2.3 (b) that $\{x_n\}$ is a Cauchy sequence and thus $x_n \to x_0$ as $n \to \infty$ for some $x_0 \in X$. Since $\|Ax_n - y\| \to 0$ as $n \to \infty$ and $A$ is a bounded linear operator, we have $Ax_0 = y$, i.e. $x_0$ is a solution of $Ax = y$. 

10) gives
\[
\Delta_p(x_m, x_l) \leq 2(\Delta_p(x^\dagger, x_l) - \Delta_p(x^\dagger, x_m)).
\]
According to the monotonicity (3.8) of $\{\Delta_p(x^\dagger, x_n)\}$, we can conclude that $\Delta_p(x_m, x_l) \to 0$ as $l, m \to \infty$. It then follows from lemma 2.3 (b) that $\{x_n\}$ is a Cauchy sequence and thus $x_n \to x_0$ as $n \to \infty$ for some $x_0 \in X$. Since $\|Ax_n - y\| \to 0$ as $n \to \infty$ and $A$ is a bounded linear operator, we have $Ax_0 = y$, i.e. $x_0$ is a solution of $Ax = y$. 

Finally we show $x_n = x^\dagger$. From the optimality condition for $x_n$ and an induction argument, we can see that $J_p(x_n) - J_p(x_0)$ is in the range of $A^*$ which implies that $(J_p(x_n) - J_p(x_0), z) = 0$ for all $z \in \mathcal{N}(A)$. This together with (3.1) gives

\[ (J_p(x_n) - J_p(x^\dagger), z) = 0, \quad z \in \mathcal{N}(A). \]

Since $J_p$ is norm-to-weak continuous, we may take $n \to \infty$ to derive that

\[ (J_p(x_n) - J_p(x^\dagger), z) = 0, \quad z \in \mathcal{N}(A). \]

Since $x_n - x^\dagger \in \mathcal{N}(A)$, we obtain $(J_p(x_n) - J_p(x^\dagger), x_n - x^\dagger) = 0$. By the strict monotonicity of $J_p$, see lemma 2.1 (a), we obtain $x_n = x^\dagger$. $\square$

**Lemma 3.2.** Let $\mathcal{X}$ and $\mathcal{Y}$ be two Banach spaces with $\mathcal{X}$ being smooth and uniformly convex, $(\alpha_n)$ be a sequence of positive numbers satisfying $\sum_{n=1}^{\infty} \alpha_n^{-1} = \infty$ and $1 < p < \infty$ and $1 \leq r < \infty$ in the definition of $\{x^\dagger_n\}$. Then the discrepancy principle (3.3) with $\tau > 1$ terminates the iteration after $n_\delta \leq \infty$ steps. If $n_\delta \geq 2$, then there hold

\[ \Delta_p(x^\dagger, x^\dagger_{n-1}) \leq \Delta_p(x^\dagger, x^\dagger_{n-1}), \quad (3.11) \]

\[ \frac{1}{\alpha_n} \|y^\delta - Ax^\dagger_n\|^r \leq \frac{\tau}{\tau - 1} (\Delta_p(x^\dagger, x^\dagger_{n-1}) - \Delta_p(x^\dagger, x^\dagger_n)) \quad (3.12) \]

for $1 \leq n < n_\delta$ and

\[ \Delta_p(x^\dagger, x^\dagger_{n_\delta}) \leq \Delta_p(x^\dagger, x^\dagger_{n_\delta-1}) + \tau^{-1} \frac{\delta'}{\alpha_{n_\delta}}. \quad (3.13) \]

**Proof.** By the similar argument for deriving (3.7), we can obtain

\[ \Delta_p(x^\dagger, x^\dagger_n) - \Delta_p(x^\dagger, x^\dagger_{n-1}) \leq \frac{1}{\alpha_n} (\xi_n^\delta, Ax^\dagger_n - y), \]

where $\xi_n^\delta \in J_p(y^\delta - Ax^\dagger_n)$. In view of property (3.6) of the mapping $J_p$ and (1.2), it follows

\[ \Delta_p(x^\dagger, x^\dagger_n) - \Delta_p(x^\dagger, x^\dagger_{n-1}) \leq \frac{1}{\alpha_n} \left( -\|y^\delta - Ax^\dagger_n\|^r + \|y^\delta - Ax^\dagger_{n-1}\|^{-1} \delta \right). \quad (3.14) \]

Since $\|y^\delta - Ax^\dagger_n\| > \tau \delta$ for $n < n_\delta$, we have for $n < n_\delta$ that

\[ \Delta_p(x^\dagger, x^\dagger_n) - \Delta_p(x^\dagger, x^\dagger_{n-1}) \leq \left( 1 - \frac{1}{\tau} \right) \frac{1}{\alpha_n} \|y^\delta - Ax^\dagger_n\|^r. \]

Since $\tau > 1$, this gives (3.11) and (3.12) immediately. In view of (3.12), we have for all $n < n_\delta$ that

\[ \sum_{j=1}^{n} \frac{1}{\alpha_j} \|y^\delta - Ax^\dagger_j\|^r \leq \frac{\tau}{\tau - 1} \Delta_p(x^\dagger, x_0). \]

By using the monotonicity (3.4) of $\{\|y^\delta - Ax^\dagger_n\|\}$, we have

\[ \|y^\delta - Ax^\dagger_n\| \sum_{j=1}^{n} \alpha_j^{-1} \leq \frac{1}{\alpha_n} \|y^\delta - Ax^\dagger_n\|^r, \quad 1 \leq n < n_\delta. \quad (3.15) \]

Using again $\|y^\delta - Ax^\dagger_n\| > \tau \delta$ for $n < n_\delta$ and $\sum_{j=1}^{n} \alpha_j^{-1} \to \infty$ as $n \to \infty$, we conclude that $n_\delta$ must be finite.

Finally, it follows from (3.14) that

\[ \Delta_p(x^\dagger, x^\dagger_{n_\delta}) - \Delta_p(x^\dagger, x^\dagger_{n_\delta-1}) \leq \frac{1}{\alpha_{n_\delta}} \delta \|y^\delta - Ax^\dagger_{n_\delta}\|^r. \]

Since $\|Ax^\dagger_{n_\delta} - y^\delta\| \leq \tau \delta$, we obtain (3.13) immediately. $\square$
**Remark 3.3.** When \( \{\alpha_n\} \) is a geometric decreasing sequence, i.e. \( \alpha_n = \alpha_0 \theta^n \) for some \( \alpha_0 > 0 \) and \( 0 < \theta < 1 \), then the method terminates after \( n_\delta = O(1 + |\log \delta|) \) steps. In fact, we now have

\[
\sum_{j=1}^{n} \alpha_j^{-1} \leq \alpha_0^{-1} (\theta^{-n} - 1)/(1 - \theta) \geq \alpha_0^{-1} \theta^{-n}.
\]

This together with (3.15) and the fact \( \|Ax_n^\delta - y^\delta\| > \tau \delta \) for \( n < n_\delta \) gives

\[
\tau \delta \theta^{-n_\delta+1} \leq C \Delta_p(x^\delta, x_0)
\]

with \( C = \alpha_0/|\tau r^{-1} (\tau - 1)| \), which implies the assertion immediately.

**Lemma 3.4.** Let \( X \) and \( Y \) be two Banach spaces with \( X \) being uniformly smooth and uniformly convex. Let \( 1 < p < \infty \) and \( 1 \leq r < \infty \) in the definition of \( \{x_n\} \) and \( \{x_n^\delta\} \). Then, for each fixed \( n \), there holds \( x_n^\delta \to x_n \) as \( y^\delta \to y \).

**Proof.** We will show this result by induction. It is trivial when \( n = 0 \) since \( x_0^\delta = x_0 \). Now we assume that \( x_{n-1}^\delta \to x_{n-1} \) and show that \( x_n^\delta \to x_n \) as \( y^\delta \to y \).

We will adapt the argument from [6]. Let \( \{y^J\} \subset Y \) be such that \( y^J \to y \) as \( J \to \infty \) and \( \{x_n^\delta, y^\delta\} \) be the minimizers of (3.2) with \( \delta \) replaced by \( \delta^J \). By the minimality of \( x_n^\delta \), we have

\[
\alpha_n \Delta_p(x_n^\delta, x_n^\delta - x_{n-1}^\delta) \leq \frac{1}{r} \|y^\delta - Ax_n^\delta\|.
\]

This implies the boundedness of \( \{\Delta_p(x_n^\delta, x_n^\delta - x_{n-1}^\delta)\} \) and hence the boundedness of \( \{\|y^\delta\|\} \). Since \( X \) is reflexive, by taking a subsequence if necessary, we may assume that \( x_n^\delta \to \bar{x}_n \) as \( J \to \infty \) for some \( \bar{x}_n \in X \). Since \( A \) is a bounded linear operator, we have \( y^\delta - Ax_n^\delta \to y - A\bar{x}_n \) as \( J \to \infty \). Recall that the uniform smoothness of \( X \) implies the continuity of \( J_p \), see lemma 2.2 (b); therefore we have from the induction hypothesis that

\[
\Delta_p(\bar{x}_n, x_{n-1}) \leq \liminf_{J \to \infty} \Delta_p(x_n^\delta, x_{n-1}^\delta)
\]

as \( J \to \infty \). Thus, by the weak lower semi-continuity of the norms in Banach spaces, we can obtain

\[
\Delta_p(\bar{x}_n, x_{n-1}) \leq \liminf_{J \to \infty} \Delta_p(x_n^\delta, x_{n-1}^\delta)
\]

and

\[
\|y - A\bar{x}_n\| \leq \liminf_{J \to \infty} \|y^\delta - Ax_n^\delta\|.
\]

Therefore, in view of the minimality of \( x_n^\delta \) and the induction hypothesis, we obtain

\[
\frac{1}{r} \|y - A\bar{x}_n\| + \alpha_n \Delta_p(\bar{x}_n, x_{n-1}) \leq \liminf_{J \to \infty} \left\{ \frac{1}{r} \|y^\delta - Ax_n^\delta\| + \alpha_n \Delta_p(x_n^\delta, x_{n-1}^\delta) \right\}
\]

\[
\leq \limsup_{J \to \infty} \left\{ \frac{1}{r} \|y^\delta - Ax_n^\delta\| + \alpha_n \Delta_p(x_n^\delta, x_{n-1}^\delta) \right\}
\]

\[
= \frac{1}{r} \|y - Ax_n\| + \alpha_n \Delta_p(x_n, x_{n-1}).
\]

By the minimality of \( x_n \) and its uniqueness, we must have \( \bar{x}_n = x_n \) and thus \( x_n^\delta \to x_n \) as \( J \to \infty \).

Next, we will show that

\[
\lim_{J \to \infty} \Delta_p(x_n^\delta, x_{n-1}^\delta) = \Delta_p(x_n, x_{n-1}).
\]
Let
\[ a := \limsup_{j \to \infty} \Delta_p(x^j_n, x^{j-1}_{n-1}) \quad \text{and} \quad b := \Delta_p(x_n, x_{n-1}). \]

According to (3.17), it suffices to show that \( a \leq b \). Assume in contrast that \( a > b \). Then, by picking a subsequence of \( \{ \delta_j \} \) if necessary, we may assume for large \( j \) that
\[ \Delta_p(x^j_n, x^{j-1}_{n-1}) \geq \limsup_{j \to \infty} \Delta_p(x^j_n, x^{j-1}_{n-1}) = \frac{a - b}{4} \tag{3.20} \]
and
\[ \frac{1}{r} \| y^{j} - Ax^{j}_{n} \|_{r} \geq \liminf_{j \to \infty} \frac{1}{r} \| y^{j} - Ax^{j}_{n} \|_{r} - \frac{\alpha_n(a - b)}{2}. \tag{3.21} \]

Therefore, by using (3.18), (3.20) and (3.21), we obtain
\[ \frac{1}{r} \| y - Ax_{n} \|_{r} + \alpha_n \Delta_p(x_n, x_{n-1}) \leq \liminf_{j \to \infty} \frac{1}{r} \| y^{j} - Ax^{j}_{n} \|_{r} + \alpha_n \limsup_{j \to \infty} \Delta_p(x^j_n, x^{j-1}_{n-1}) - \alpha_n(a - b) \]
for large \( j \). Since \( \alpha_n(a - b) > 0 \), we have from the induction hypothesis that
\[ \frac{1}{r} \| y^{j} - Ax^{j}_{n} \|_{r} + \alpha_n \Delta_p(x_n, x_{n-1}) < \frac{1}{r} \| y^{j} - Ax^{j}_{n} \|_{r} + \alpha_n \Delta_p(x^j_n, x^{j-1}_{n-1}) \]
for large \( j \), which contradicts the minimality of \( x^j_n \). We therefore obtain (3.19).

In view of (3.16), we can conclude from (3.19) that \( \| x^{j}_{n} \| \to \| x_{n} \| \) as \( j \to \infty \). Since \( X \) is uniformly convex and \( x^{j}_{n} \to x_{n} \), it follows from lemma 2.1 (b) that \( \| x^{j}_{n} - x_{n} \| \to 0 \) as \( j \to \infty \).

The above argument shows that for any sequence \( \{ y^{j} \} \) converging to \( y \), the sequence \( \{ x^{j}_{n} \} \) always has a subsequence that converges to \( x_{n} \). Therefore, \( x^{j}_{n} \to x_{n} \) as \( y^{j} \to y \). \( \square \)

**Theorem 3.5.** Let \( X \) and \( Y \) be two Banach spaces with \( X \) being uniformly smooth and uniformly convex, \( \{ \alpha_n \} \) be a sequence of positive numbers satisfying \( \sum_{n=1}^{\infty} \alpha_n^{-1} = \infty \) and \( \alpha_n \leq c_0 \alpha_{n+1} \) for all \( n \) for some constant \( c_0 > 0 \) and \( 1 < p < \infty \) and \( 1 \leq r < \infty \) in the definition of the nonstationary iterated Tikhonov regularization in Banach spaces. Then the discrepancy principle (3.3) with \( \tau > 1 \) terminates the method after \( n_{\tau} < \infty \) steps. Moreover, \( x^{j}_{n} \) converges to the solution \( x^+ \) of (1.1) as \( \delta \to 0 \).

**Proof.** It remains to show the convergence of \( x^{j}_{n_{\tau}} \) to \( x^+ \) since other parts have been proved in lemma 3.2.

Assume first that \( \{ y^{j} \} \) is a sequence satisfying \( \| y^{j} - y \| \leq \delta_j \) with \( \delta_j \to 0 \) such that \( n_{\delta_j} \to n_{\tau} \) as \( j \to \infty \) for some finite integer \( n_{\tau} \). We may assume \( n_{\delta_j} = n_{\tau} \) for all \( j \). From the definition of \( n_{\tau} = n_{\delta_j} \), we have
\[ \| Ax^{j}_{n_{\tau}} - y^{j} \| \leq \tau \delta_j. \]

Since lemma 3.4 implies \( x^{j}_{n_{\tau}} \to x_{n_{\tau}} \), by letting \( j \to \infty \) we obtain \( Ax_{n_{\tau}} = y \). This together with the definition of \( x_{n} \) implies \( x_n = x_{n_{\tau}} \) for all \( n \geq n_{\tau} \). Since lemma 3.1 implies that \( x_n \to x^+ \) as \( n \to \infty \), we must have \( x_{n_{\tau}} = x^+ \). Consequently \( x^{j}_{n_{\tau}} \to x^+ \) as \( j \to \infty \).

Assume next that \( \{ y^{j} \} \) is a sequence satisfying \( \| y^{j} - y \| \leq \delta_j \) with \( \delta_j \to 0 \) such that \( n_{\delta_j} := n_{\delta_j} \to \infty \) as \( j \to \infty \). We first show that \( \Delta_p(x^+, x^{j}_{n_{\tau}-1}) \to 0 \) as \( j \to \infty \). Let \( \epsilon > 0 \) be an arbitrary number. Since \( x_n \to x^+ \), we may pick an integer \( n(\epsilon) \) such that
Thus \( J_p(x^*, x_{n(e)}) < \epsilon / 2 \). Since \( x_{n(e)}^j \to x_{n(e)} \) and the uniform smoothness of \( \mathcal{X} \) implies that \( J_p \) is continuous, see lemma 2.2 (b), we can take an integer \( j(\epsilon) \) such that \( n_j \geq n(\epsilon) \) and \( | \Delta_p(x^*, x_{n(e)}^j) - \Delta_p(x^*, x_{n(e)}) | < \epsilon / 2 \) for all \( j \geq j(\epsilon) \). Consequently, by using (3.11) it follows that

\[
\Delta_p(x^*, x_{n-2}^j) \leq \Delta_p(x^*, x_{n(e)}^j) \leq \Delta_p(x^*, x_{n(e)}) + | \Delta_p(x^*, x_{n(e)}^j) - \Delta_p(x^*, x_{n(e)}) | < \epsilon
\]

for all \( j \geq j(\epsilon) \). Since \( \epsilon > 0 \) is arbitrary, we must have \( \Delta_p(x^*, x_{n-2}^j) \to 0 \) as \( j \to \infty \). This together with (3.11) implies also that \( \Delta_p(x^*, x_{n-1}^j) \to 0 \) as \( j \to \infty \).

In view of (3.12), we have

\[
\frac{1}{\alpha_{n_j-1}} \| y^j - Ax_{n_j-1}^j \|^2 \leq \frac{\tau}{\tau - 1} \Delta_p(x^*, x_{n_j-2}^j).
\]

Since \( \| y^j - Ax_{n_j-2}^j \| > \tau \delta_j \), it gives

\[
\delta_j \frac{1}{\alpha_{n_j-1}} \leq \frac{1}{\tau - 1} \Delta_p(x^*, x_{n_j-2}^j).
\]

This implies that \( \delta_j / \alpha_{n_j-1} \to 0 \) as \( j \to \infty \). Since \( \alpha_n \leq c_0 \alpha_{n+1} \), we have \( \delta_j / \alpha_{n_j} \to 0 \) as \( j \to \infty \). In view of (3.13), we therefore obtain \( \Delta_p(x^*, x_{n_j}^j) \to 0 \) as \( j \to \infty \). Consequently, by using lemma 2.3 (a), we obtain \( x_{n_j}^j \to x^* \) as \( j \to \infty \).

\[\square\]

**Remark 3.6.** The conditions on \( \{\alpha_n\} \) in theorem 3.5 are standard and cover the well-known cases that \( \alpha_n = 1 \) for all \( n \) and that \( \{\alpha_n\} \) is a geometric decreasing sequence.

**Remark 3.7.** The result in theorem 3.5 requires \( \mathcal{X} \) to be uniformly convex and uniformly smooth. However, it does not require any geometric information on the Banach space \( \mathcal{Y} \).

**Remark 3.8.** We may drop the condition \( \alpha_n \leq c_0 \alpha_{n+1} \) on \( \{\alpha_n\} \) in the convergence analysis if \( 1 < r < \infty \) and if we replace the discrepancy principle (3.3) by the variant that defines \( \hat{n}_3 \) to be the largest integer such that \( \| Ax_n^j - y^j \| \geq \tau \delta \), i.e.

\[
\| Ax_{n+1}^j - y^j \| < \tau \delta \equiv \| Ax_{n}^j - y^j \|, \quad 0 \leq n \leq \hat{n}_3,
\]

where \( \tau > 1 \) is a given number. The proof of lemma 3.2 can be used without change to show that \( \hat{n}_3 < \infty \) and

\[
\Delta_p(x^*, x_{n+1}^j) \leq \Delta_p(x^*, x_{n-1}^j), \quad 1 \leq n \leq \hat{n}_3.
\]

To see the convergence of \( x_{n+1}^j \) to \( x^* \), we need to consider two cases as in the proof of theorem 3.5. For the case that \( \{y^j\} \) is a sequence satisfying \( \| y^j - y \| \leq \delta_j \) with \( \delta_j \to 0 \) such that \( \hat{n}_3_j \to \infty \) as \( j \to \infty \), we can follow the derivation of \( \Delta_p(x^*, x_{n_j}^j) \to 0 \) to show that \( \Delta_p(x^*, x_{n_j}^j) \to 0 \) as \( j \to \infty \). We only need to consider the case that \( \{y^j\} \) is a sequence satisfying \( \| y^j - y \| \leq \delta_j \) with \( \delta_j \to 0 \) such that \( \hat{n}_3_j \to n_0 \) as \( j \to \infty \) for some finite integer \( n_0 \). We may assume \( \hat{n}_3_j = n_0 \) for all \( j \). Now we only have \( \| Ax_{n_j}^j - y^j \| \geq \tau \delta_j \). However, we have \( \| Ax_{n_j}^j - y^j \| < \tau \delta_j \). By taking \( j \to \infty \), we obtain \( Ax_{n_j} \to y \). According to the definition of \( x_n \) we have \( x_n = x_{n+1} \) for all \( n \geq n_0 + 1 \). It then follows from lemma 3.1 that \( x_{n_0+1} = x^* \). In view of the optimality condition for \( x_{n_0+1} \) and the fact that \( J_r(0) = 0 \) for \( 1 < r < \infty \), we have

\[
\alpha_{n_0+1} \left( J_p(x_{n_0+1}) - J_p(x_{n_0}) \right) \in A^* J_f(y - Ax_{n+1})
\]

\[
= A^* J_f(0) = 0.
\]

Thus \( J_p(x_{n_0+1}) - J_p(x_{n_0}) = 0 \). By using the strict monotonicity of \( J_p \), see lemma 2.1 (a), we obtain \( x_{n_0+1} = x_{n_0} \). Consequently, \( x_{n_0} = x^* \). Since \( x_{n_0} \to x_{n_0}^j \), we thus obtain \( x_{n_0}^j \to x^* \) as \( j \to \infty \).
4. Numerical examples

We consider the integral equation of the form

$$Ax(s) := \int_0^1 K(s, t)x(t) \, dt = y(s) \quad \text{on} \ [0, 1],$$

(4.1)

where $K(s, t)$ is a continuous function defined on $[0, 1] \times [0, 1]$. It is easy to see, for any $1 \leq p \leq \infty$, that $A$ is a bounded linear map from $L^p[0, 1]$ to $C[0, 1]$ and $\|Ax\|_{L^\infty[0, 1]} \leq M \|x\|_{L^p[0, 1]}$ with

$$M := \max_{x \in [0, 1]} \left(\int_0^1 |K(s, t)|^p \, dt\right)^{1/p}, \quad \frac{1}{p} + \frac{1}{p'} = 1.$$ 

Thus $A : L^p[0, 1] \to L^r[0, 1]$ is a compact operator for any $1 \leq r < \infty$.

We will take $X = L^p[0, 1]$ and $Y = L^r[0, 1]$ with $1 < p < \infty$ and $1 < r < \infty$. The adjoint of $A$ is an operator $A^* : L^r[0, 1] \to L^p[0, 1]$ given by

$$A^*y(s) = \int_0^1 K(t, s)y(t) \, dt.$$ 

Our question is to find the solution of (4.1) by using some noisy data $y^\delta$ of $y$. By choosing an initial guess $x_0$, our method in fact defines each $x_n^\delta$ as the minimizer of the functional

$$\frac{1}{r} \int_0^1 |Ax - y^\delta|^r + \alpha_n \left(\frac{1}{p} \int_0^1 |x|^p - \int_0^1 J_p(x^\delta_{n-1})(x - x^\delta_{n-1})\right)$$

(4.2)

over $L^p[0, 1]$, where $J_p$ is the duality mapping in $L^p[0, 1]$ and $J_p(x)$, for each function $x \in L^p[0, 1]$, has the pointwise expression

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

For our numerical simulations, we divide $[0, 1]$ into $N$ subintervals of equal length with nodal points

$$t_j = \frac{j}{N}, \quad j = 0, 1, \ldots, N.$$ 

We identify any function by the vector whose components are its values at $t_j$, i.e. $x \sim (x(t_j))_{j=0}^N$. We also approximate integrals by the trapezoidal rule. In order to produce the next iterate, instead of finding the minimizer of (4.2) directly, we solve a neighboring minimization problem by replacing any integral of the form $\int_0^1 |f|^q$ with $1 < q < 2$ by the smooth one $\int_0^1 (f^2 + \varepsilon^2)^{q/2}$ for a sufficiently small $\varepsilon > 0$; we take $\varepsilon = 10^{-8}$. We then solve the corresponding minimization problem by a (quasi)-Newton method.

**Example 4.1.** We consider equation (4.1) with

$$K(s, t) = \begin{cases} 40s(1-t), & s \leq t \\ 40(1-s), & s \geq t. \end{cases}$$

(4.3)

We assume that the solution is given by

$$x^\dagger(t) = \begin{cases} 1, & [0.225, 0.275] \text{ and } [0.725, 0.775] \\ 2, & [0.475, 0.525] \\ 0, & \text{elsewhere}. \end{cases}$$

Let $y := Ax^\dagger$ which are the exact data. We add random Gaussian noise to $y$ to get the noisy data $y^\delta$ satisfying

$$\|y - y^\delta\|_{L^2[0,1]} = \delta$$

for a specified noise level $\delta$ and use $y^\delta$ in our method to recover $x^\dagger$. 

Figure 1. Reconstruction results for example 4.1 with \( n_\delta \) being the number of iterations. (a) \( n_\delta = 16 \); (b) \( n_\delta = 16 \); (c) \( n_\delta = 24 \).

In figure 1, we present the reconstruction results via the nonstationary iterated Tikhonov regularization, where we take \( N = 400 \), \( \delta = 0.5 \times 10^{-3} \), \( \alpha_n = 2^{-n} \) and \( \tau = 1.01 \). Figure 1(a) reports the result via the 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 special feature of the exact solution \( x^* \). 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 zero part is much better; however, we have to pay the price that the reconstruction is oscillatory on the non-zero part which is well known for the \( L^p \)-regularization with \( p \geq 1 \) but close to 1.

In order to get rid of this notorious effect on nonzero part, we may need to consider the TV-like regularization [16]. We may extend our methods to more general situations where the Bregman distance is induced by a convex function \( f : \mathcal{X} \to \mathbb{R} \) which, for simplicity, is assumed to be Fréchet differentiable at every \( x \in \mathcal{X} \) with gradient \( f'(x) : \mathcal{X} \to \mathcal{X}^* \). The corresponding Bregman distance is

\[
D_f(x, \bar{x}) := f(x) - f(\bar{x}) - \langle f'(\bar{x}), x - \bar{x} \rangle.
\]

Then, once a current iterate \( x^\delta_{n-1} \) is available, we define the next iterate \( x^\delta_n \) to be the minimizer of the minimization problem

\[
\min_{x \in \mathcal{X}} \left\{ \frac{1}{r} \| Ax - y^\delta \|_r^r + \alpha_n D_f(x, x^\delta_{n-1}) \right\}, \tag{4.4}
\]
where \( \{\alpha_n\} \) is a preassigned sequence of positive numbers. The common choices of \( f \) are
\[
f(x) = \int_0^1 |x'(t)|^p \, dt \quad \text{with} \quad 1 \leq p < \infty,
\]
which, however, are not well defined on the whole space. In our calculation, we take the discrete version
\[
f(x) = N^{p-1} \sum_{j=1}^N |x(t_j) - x(t_{j-1})|^p, \quad 1 < p < \infty.
\]
We take \( p = 1.1 \) and the result is presented in figure 1(c). Our theory does not apply directly to this case; we hope to develop a convergence theory in a forthcoming paper.

**Example 4.2.** Next we give a reconstruction result for the case that the data contain non-Gaussian noise. Thus there may exist data points that are highly inconsistent with other data points. Such data points are called outliers, which may arise from procedural measurement error in many practical applications including image processing.

We consider again equation (4.1) with \( K(s, t) \) given by (4.3). We assume that the solution to be constructed is
\[
x^\dagger(t) = \frac{1}{1 + 80(t - 0.5)^2} - \frac{1}{21}.
\]
Let \( y := Ax^\dagger \) and \( y^\delta \) be the noisy data of \( y \). In figure 2, we present the reconstruction results for \( x^\dagger \) from \( y^\delta \) by using our method with \( \alpha_n = 2^{-n}, \tau = 1.01 \) and \( N = 400 \). Figures 2(a) and (d) present the plots of the noisy data; the one in (a) contains not only Gaussian noise but also a few outliers, while the one in (d) contains only Gaussian noise. Figures 2(b) and (e)
present the reconstruction results by the nonstationary iterated Tikhonov regularization with $\mathcal{X} = L^2[0, 1]$ and $\mathcal{Y} = L^2[0, 1]$. 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 $\mathcal{X} = L^2[0, 1]$ and $\mathcal{Y} = L^1[0, 1]$ which corresponds to the $L^1$-data fitting. It can be seen that the method is robust enough to prevent the affection from outliers. Therefore, $L^1$-data fitting can be used to deal not only with Gaussian noise but also non-Gaussian noise efficiently.

5. Conclusion and questions

By making use of duality mappings and the Bregman distance, we extended the nonstationary iterated Tikhonov regularization from the Hilbert space setting to the Banach space setting. We obtained the convergence of the method when a discrepancy principle is used to terminate the iteration. Our convergence result requires the pre-image space to be uniformly smooth and uniformly convex. However, it does not require any geometric information on the image space. We also presented numerical results which indicate that the method in Banach spaces can produce better reconstructions in some applications.

The following issues might be interesting for further investigations.

(a) Can we develop fast and efficient algorithms to solve the convex minimization problem (3.2)? This would make our method more favorable since (3.2) is the main ingredient.

(b) Our result only provides convergence of the method. Is it possible to derive the rate of convergence when the source condition

$$J_p(x^\dagger) - J_p(x_0) \in R(A^*)$$

is satisfied? More generally, can we establish a convergence rate result under the source conditions formulated as variational inequalities introduced in [12]?

(c) If we replace the penalty term in (3.2) by the Bregman distance from a general convex function like the one in (4.4), can we develop a general convergence result for the corresponding method?

(d) How can we extend our method to solve nonlinear ill-posed inverse problems in Banach spaces?

Acknowledgment

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

References


