Iterative regularization for elliptic inverse problems

A.A. Khan\textsuperscript{a,}\textsuperscript{*}, B.D. Rouhani\textsuperscript{b}

\textsuperscript{a} Department of Mathematics, University of Wisconsin-Barron County, 1800 College Drive, Rice Lake, WI 54868, United States
\textsuperscript{b} Department of Mathematical Sciences, University of Texas at El Paso, El Paso, TX 79968, United States

Received 10 October 2006; received in revised form 23 January 2007; accepted 20 February 2007

Dedicated to the 70th birthday of Prof. Helmut Neunzert

Abstract

Elliptic inverse problems can be formulated using coefficient-dependent energy least-squares functionals, resulting in a smooth, convex objective functional. A variational inequality emerges as a necessary and sufficient optimality condition. The principle of iterative regularization, when coupled with the auxiliary problem principle, results in a strongly convergent scheme for the solution of elliptic inverse problems.

\textcopyright{} 2007 Elsevier Ltd. All rights reserved.

Keywords: Inverse problem; Parameter identification; Variational inequality; Regularization; Auxiliary problem principle; Iterative methods

1. Introduction

The problem of identifying parameters in elliptic boundary value problems has applications in several fields such as groundwater management, identifying cracks, modeling of car wind-shields, image processing and many others (see [5,6,10]). Therefore, it is not surprising that a plentiful literature is devoted to the theoretical and numerical investigation of these problems. One of the most commonly studied parameter identification problems arises in the context of the following elliptic boundary value problem (BVP):

\begin{align}
-\nabla \cdot (a \nabla u) &= f \quad \text{in } \Omega, \\
u &= 0 \quad \text{on } \partial \Omega.
\end{align}

The above BVP models several interesting real-world problems and has been studied in great detail (see [11]). For example, $u = u(x)$ may represent the steady-state temperature at a given point $x$ of a body; then $a$ would be a variable thermal conductivity coefficient and $f$ the external heat source. The system (1) also models underground steady state aquifers in which the coefficient $a$ is the aquifer transmissivity coefficient, $u$ is the hydraulic head, and $f$ is the recharge. In either context, the \textit{direct problem} is to find $u$ given the coefficient $a$ and the source term $f$. On the other hand the inverse problem is to estimate the coefficient $a$, given some measurement of $u$.

\textsuperscript{*} Corresponding author.

E-mail addresses: akhan@uwc.edu (A.A. Khan), behzad@math.utep.edu (B.D. Rouhani).
In the context of (1), there have been mainly two approaches for attacking the corresponding inverse problem. The first approach reformulates the inverse problem as an optimization problem, and then employs some suitable method for its solution. The second approach treats (1) as a hyperbolic partial differential equation in \( a \). Furthermore, the approach of reformulating (1) as an optimization problem is divided into two possibilities, namely either formulating the problem as an unconstrained optimization problem, or treating it as a constrained optimization problem in which the PDE itself is the constraint. Among the optimization-based techniques the output least-squares method is among the most widely investigated methods (see [19]). The output least-squares approach minimizes the functional
\[
J_1(a) = \| u(a) - z \|^2, \tag{2}
\]
where \( z \) is the data (the measurement of \( u \)) and \( u(a) \) solves the variational form of (1) given by
\[
\int_\Omega a \nabla u \cdot \nabla v = \int_\Omega f v \quad \forall v \in H^1_0(\Omega). \tag{3}
\]
However, a deficiency of the output least squares functional is that it often fails to be convex. In [13], we studied common features of elliptic inverse problems by establishing an abstract framework based on the modified output least-squares approach. The main contribution of our approach in [13] was a proposed objective function which used a coefficient-dependent energy norm. This objective functional turned out to be smooth and convex, allowing us to use several powerful existing results for the theoretical and numerical treatment of inverse problems. Recall that for (1), the modified output least-squares approach minimizes the functional
\[
J_1(a) = \int a \nabla (u(a) - z) \cdot \nabla (u(a) - z), \tag{4}
\]
where \( z \) is the data (the measurement of \( u \)) and \( u(a) \) solves (3). We note that the above functional was studied, independently, in [18,24].

The main objective of this work is to study some possibilities of exploiting the convexity of the objective functional proposed in [13] at an abstract level. In particular, we aim to employ the principle of iterative regularization (PIR) for the numerical solvability of elliptic inverse problems (see [4]). Roughly speaking, in the PIR approach the regularization parameter is modified at each iteration, which is in contrast with the common practice for parameter identification of using a fixed regularization parameter throughout the minimization process. The main numerical scheme is designed and studied in context of a general variational inequality.

The paper is divided into four sections. Section 2 opens up by describing the abstract framework for elliptic inverse problems. In this section, we formulate a minimization problem and its regularized variant for the parameter identification. In Section 3 we propose an algorithm based on the PIR and present its convergence analysis. The development of this section is based on a general variational inequality which recovers the parameter identification problem. In Section 4, we discuss some computational results. The paper concludes with some remarks about the approach.

2. Problem formulation and regularization

Let \( X \) and \( Y \) be two real Hilbert spaces, let \( A \) be a nonempty closed and convex subset of \( X \), let \( T : X \times Y \times Y \to \mathbb{R} \) be a trilinear form, and let \( m : Y \to \mathbb{R} \) be a bounded linear functional. Let the trilinear form \( T(\cdot, \cdot, \cdot) \) be equipped with the following features:
\[
\begin{align*}
T(a, u, u) & \geq \alpha \| u \|_Y^3 \quad \forall a \in A, \forall u \in Y \tag{5a} \\
T(a, u, v) & \leq \beta \| a \|_X \| u \|_Y \| v \|_Y \quad \forall a \in X, \forall u \in Y, \forall v \in Y \tag{5b} \\
T(a, u, v) & = T(a, v, u) \quad \forall a \in X, \forall u \in Y, \forall v \in Y. \tag{5c}
\end{align*}
\]
Consider the following variational problem: For a fixed but arbitrary \( a \in A \), find \( u \in Y \) such that for all \( v \in Y \) the following equation holds:
\[
T(a, u, v) = m(v). \tag{6}
\]
Notice that under the standard boundedness hypothesis on the set \( \mathcal{A} \), the Riesz representation theorem ensures the existence of a unique solution of the variational problem (6). We therefore define the solution operator \( F : \mathcal{A} \rightarrow Y \) by the condition that \( u = F(a) \) is the solution to (6).

In this setting, we focus on the following inverse problem associated with the direct problem (6): Given some measurement of \( u \), say \( z \), estimate the coefficient \( a \) which together with \( u \) makes (6) true.

In applications, the data \( z \) are obtained by means of a finite set of measurements, and then interpolated in some suitable manner. We will assume that \( z \in V \).

We consider the following minimization problem to estimate the coefficient \( a \): Given a measurement \( z \in V \) of \( u \), find \( a^* \in A \subset \text{int}(\mathcal{A}) \) (\( \text{int}(\mathcal{A}) \) is the interior of \( \mathcal{A} \)) by solving

\[
\min_{a \in A} J(a) \quad \text{where} \quad J(a) = \frac{1}{2} T(a, F(a) - z, F(a) - z). \quad (7)
\]

We assume that \( A \) is closed and convex. We shall denote the set of all solutions of (7) by \( S(J) \). The functional \( J \) was introduced, from an abstract point of view, in [13] and was motivated by two independent proposals of Chen and Zou [7] (see Zou [24]) and Knowles [18].

To discuss the differentiability of \( J \), we analyze the differentiability of the solution operator \( F \). For this we recall the following result.

**Lemma 2.1** ([13]). For each \( a \) in the interior of \( \mathcal{A} \), \( F \) is infinitely differentiable at \( a \). Given \( u = F(a) \), the first derivative \( \delta u = DF(a)\delta a \) is the unique solution to the variational equation

\[
T(a, \delta u, v) = -T(\delta a, u, v) \quad \text{for all} \quad v \in Y,
\]

and \( \delta^2 u = D^2 F(a)(\delta a_1, \delta a_2) \) is the unique solution to the variational equation

\[
T(a, \delta^2 u, v) = -T(\delta a_2, DF(a)\delta a_1, v) - T(\delta a_1, DF(a)\delta a_2, v) \quad \text{for all} \quad v \in Y.
\]

The above derivatives can be used to compute derivatives of \( J \). In fact, the first derivative of \( J \) is readily obtained by using the chain rule:

\[
D\mathcal{J}(a)\delta a = \frac{1}{2} T(\delta a, F(a) - z, F(a) - z) + T(a, DF(a)\delta a, F(a) - z).
\]

Let us now compute the second derivative. By (8), we have

\[
T(a, DF(a)\delta a, F(a) - z) = -T(\delta a, F(a), F(a) - z),
\]

and hence

\[
D\mathcal{J}(a)\delta a = \frac{1}{2} T(\delta a, F(a) - z, F(a) - z) - T(\delta a, F(a), F(a) - z)
\]

\[
= -\frac{1}{2} T(\delta a, F(a) + z, F(a) - z).
\]

It now follows that

\[
D^2\mathcal{J}(a)(\delta a, \delta a) = -\frac{1}{2} T(\delta a, DF(a)\delta a, F(a) - z) - \frac{1}{2} T(\delta a, F(a) + z, DF(a)\delta a)
\]

\[
= -T(\delta a, F(a), DF(a)\delta a)
\]

\[
= T(a, DF(a)\delta a, DF(a)\delta a).
\]

In the last step, we applied (8) again. We notice, in particular, that the following inequality holds for all \( a \) in the interior of \( \mathcal{A} \):

\[
D^2\mathcal{J}(a)(\delta a, \delta a) \geq \alpha\|DF(a)\delta a\|^2.
\]

Thus \( \mathcal{J} \) is a convex functional over \( A \subset \text{int}(\mathcal{A}) \).

It is known that (7) is an ill-posed problem, and hence some regularization is needed for its stabilization. We introduce functionals \( R, \phi : X \rightarrow \mathbb{R} \) and a sequence \( \{\epsilon_n\} \) of strictly decreasing positive reals, and consider the
regularized variant of (7) as follows: Find $a_{\epsilon_n} \in A$ by solving

$$
\min_{a \in A} \mathcal{J}_{\epsilon_n}(a) \quad \text{where} \quad \mathcal{J}_{\epsilon_n} = \frac{1}{2} T(a, F(a) - z, F(a) - z) + \epsilon_n R(a) + \eta \phi(a), \quad \eta \geq 0.
$$

(9)

We need to impose conditions on $R$ and $\phi$ which ensure the solvability, uniqueness and stability of (9). For example, it suffices to assume that $R$ is Gateaux differentiable and strongly convex. One theoretical advantage of the regularization operator $R$ is that it ensures the uniqueness. We assume that $\phi$ is convex, proper and lower semicontinuous. We do not assume that $\phi$ is differentiable. The functional $\phi(\cdot)$ corresponds to total variation-type regularization. In recent years, total-variation regularization has been very successful in identifying discontinuous or rapidly varying coefficients. We remark that this kind of double regularization for a linear inverse problem was given in [15], where $\eta$ played the role of an actual regularizing parameter and $\epsilon_n = \epsilon > 0$ was chosen sufficiently small. However, in this work we are interested in the case when $\epsilon_n \to 0$. When $\eta > 0$ is held fixed, we view the present approach as applying the PIR to the problem with objective functional $\mathcal{J}(\cdot) + \eta \phi(\cdot)$, whereas $\eta = 0$ means that we only regularize through $R$. The following result summarizes some useful features of the regularized problem. Although, there are many choices for $R$, for simplicity, we assume that $R(x) = \frac{1}{2} \|x\|^2$. See [14,20] for more details.

**Lemma 2.2.** Let $S(\mathcal{J}) \neq \emptyset$ and let $x_{\epsilon_n}$ be the unique solution of (9) for $n \in \mathbb{N}$. Then the solution set $S(\mathcal{J})$ is convex and closed. Moreover, the sequence $\{x_{\epsilon_n}\}_{n=1}^{\infty}$ is bounded and converges strongly to the minimum norm element $x^*$ of $S(\mathcal{J})$. Finally, the following estimate holds:

$$
\|x_{\epsilon_n} - x_{\epsilon_{n+1}}\| \leq K \frac{\epsilon_n - \epsilon_{n+1}}{\epsilon_{n+1}} \quad (K \text{ is a positive constant}).
$$

(10)

**Proof.** The proof of the above result can be extracted from [4], where the stated properties are shown to be true for a monotone variational inequality. \qed

We conclude this section by a remark concerning a variant of the above result for the output least-squares approach. Notice that in the abstract setting, the output least-squares functional takes the following form:

$$
J(a) = \frac{1}{2} \|F(a) - z\|^2.
$$

(11)

It is known (see [17]) that the one-sided directional derivative $J'$ of $J$ satisfies

$$
(J'(a) - J'(b), a - b) \geq -\kappa \|a - b\|^2 \quad \text{for all} \quad a, b \in A
$$

where $\kappa = 2\beta^2 \alpha^{-4} \|m\|(\|m\| + \alpha \|z\|)$. Therefore, to obtain a result similar to Lemma 2.2 for (7), with $\mathcal{J}$ replaced by $J$, we can consider a sequence of regularizing parameters $\{\tilde{\epsilon}_n\}$ such that it is strictly decreasing and converges to some $\tilde{\epsilon} > \kappa$. This, however, shows that we might not be free to bring the regularized functional $J + \tilde{\epsilon}_n R$, sufficiently close to the original functional $J$, and, depending on the size of $\kappa$, we might have to choose a bigger regularization parameter $\tilde{\epsilon}$. This may lead to over-regularization of the minimization problem.

3. Iterative regularization

Let $\mathcal{H}$ be a real Hilbert space and let $\mathcal{H}^*$ be the topological dual of $\mathcal{H}$, which will be identified with $\mathcal{H}$. Let $\mathcal{K}$ be a nonempty closed and convex subset of $\mathcal{H}$, let $\mathcal{F} : D(\mathcal{F}) \to \mathcal{H}$ ($D(\mathcal{F})$ is the domain of $\mathcal{F}$) be a given operator, and let $\psi : \mathcal{H} \to \mathbb{R}$ be a given functional. We consider the following variational inequality: Find $x \in \mathcal{K}$ such that

$$
\langle \mathcal{F}(x), z - x \rangle \geq \psi(x) - \psi(z), \quad \text{for all} \quad z \in \mathcal{K}.
$$

(12)

The above variational inequality is more general than the one that actually emerges from (7). Indeed, if $\mathcal{F}(\cdot) = J'(\cdot)$, the Gateaux derivative of $J$, $A = \mathcal{K}$, and $\psi$ is the indicator functional of $\mathcal{K}$, then (12) is a necessary and sufficient optimality condition for (7).

We also consider the following regularized variational inequality: For a fixed but arbitrary $n \in \mathbb{N}$, and for $\epsilon_n > 0$, find $x_{\epsilon_n} \in \mathcal{K}$ such that

$$
\langle \mathcal{F}(x_{\epsilon_n}) + \epsilon_n x_{\epsilon_n}, z - x_{\epsilon_n} \rangle \geq \psi(x_{\epsilon_n}) - \psi(z) \quad \text{for all} \quad z \in \mathcal{K}.
$$

(13)
Our primary goal here is to develop an iterative scheme for (12) by using its regularized version (13). For this, let $\mathcal{A} : \mathcal{H} \to \mathbb{R}$ be a Gateaux differentiable functional, and let $\mathcal{A}'$ be its Gateaux derivative. Let $\{\alpha_n\}_{n=1}^{\infty}$ be a sequence of positive reals, and let $\{\epsilon_n\}_{n=1}^{\infty}$ be a decreasing sequence of positive reals such that $\epsilon_n \to 0$ as $n \to \infty$. For an approximate solution of (12), we consider:

**Algorithm 1.**

(i) At $k = 0$ start with $x^0, \epsilon^0$, and $\alpha^0$.

(ii) At step $k = n$, compute $x^{n+1} \in \mathcal{K}$ by solving the minimization problem

$$
\min_{x \in \mathcal{K}} \{\mathcal{A}(x) + \langle \alpha_n(\mathcal{F}_n(x^n) + \epsilon_n x^n) - \mathcal{A}'(x^n), x \rangle + \alpha_n \psi(x^n)\}.
$$

(iii) Stop if $\|x^{n+1} - x^n\|$ is below some desired threshold. Otherwise, make $n + 1 = n$, and go (ii).

Before presenting the convergence analysis for the above algorithm, we have a few remarks to make. The above algorithm is based on a coupling of the auxiliary problem principle (APP) and the principle of iterative regularization (PIR). As already mentioned, PIR was introduced by A.B. Bakushinskii in connection with variational inequalities. An important extension of this approach is presented by Alber [1]. The APP was introduced by G. Cohen [8] (see [21, 22] and the references therein) to unify various existing optimization methods. It is now known that if the auxiliary functional $\mathcal{A}(\cdot)$ is allowed to change at each iteration, then Newton-type methods and proximal point methods can be recovered from the APP (see [3,12,16,23]). In particular, if $\psi \equiv 0$ and $\mathcal{A}(\cdot) = \frac{1}{2} \| \cdot \|^2$, then (14) is equivalent to the projection method:

$$
x^{n+1} = P_{\mathcal{K}}[x^n - \alpha_n(\mathcal{F}_n(x^n) + \epsilon_n x^n)].
$$

The following result gives the convergence of the above algorithm.

**Theorem 3.1.** Let $S(\mathcal{J}) \neq \emptyset$, let $\mathcal{F} : D(\mathcal{F}) \to \mathcal{H}$ be single-valued, monotone and Lipschitz continuous, let $\psi : \mathcal{H} \to \mathbb{R}$ be proper, convex and lower semicontinuous. Assume that $\mathcal{A} : \mathcal{H} \to \mathbb{R}$ be proper, strongly convex and Gateaux differentiable with strongly monotone and Lipschitz continuous Gateaux derivative $\mathcal{A}'$. For the approximation $\mathcal{F}_n$ of $\mathcal{F}$, assume that there exists $\{\delta_n\}$ with $\delta_n \geq 0$, such that

$$
\|\mathcal{F}_n(x) - \mathcal{F}(x)\| \leq c\delta_n(1 + \|x\|), \quad \text{for all } x \in S(\mathcal{J}), \quad (c \text{ is a constant}).
$$

Assume that the sequences $\{\epsilon_n\}, \{\alpha_n\}$ and $\{\delta_n\}$ satisfy the following conditions:

$$
\sum_{n=0}^{\infty} \epsilon_n \alpha_n = \infty; \quad \sum_{n=0}^{\infty} \left(\alpha_n^2 + \delta_n^2\right) < \infty; \quad \left\{\frac{\alpha_n}{\epsilon_n} + \delta_n \epsilon_n\right\} \to 0; \quad \sum_{n=0}^{\infty} \frac{|\epsilon_n - \epsilon_{n-1}|}{\epsilon_n^2 \alpha_n} < \infty.
$$

Then:

(a) For each $n \in \mathbb{N}$, (14) is uniquely solvable.

(b) The sequence $\{x^n\}_{n=1}^{\infty}$, generated by Algorithm 1, is bounded.

(c) If there is a constant $\ell$ such that, for $n \in \mathbb{N}$, the inequality

$$
|\psi(x^{n+1}) - \psi(x^n)| \leq \ell\|x^{n+1} - x^n\|
$$

holds, then there exists a constant $\delta_0$ such that

$$
\|x^{n+1} - x^n\| \leq \delta_0 \alpha_n.
$$

(d) If the functional $\mathcal{A}$ is defined by

$$
\mathcal{A}(x) = \frac{1}{2} a(x, x), \quad \text{for all } x \in \mathcal{H},
$$

where $a : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ is a symmetric, coercive and continuous bilinear form, then the sequence $\{x^n\}_{n=1}^{\infty}$ converges strongly to the minimal norm element $x^*$ of $S(\mathcal{J})$.

**Proof.** (a) The minimization problem (14) is equivalent to VI: Given $x^n \in \mathcal{K}$, find $x \in \mathcal{K}$ so that

$$
\langle \mathcal{A}'(x) - \mathcal{A}'(x^n) + \alpha_n(\mathcal{F}_n(x^n) + \epsilon_n x^n), z - x \rangle \geq \alpha_n(\psi(x) - \psi(z)) \quad \forall z \in \mathcal{K}.
$$

(19)

In view of the conditions imposed on $\mathcal{A}$ and $\psi$, (19) has a unique solution $x^{n+1}$. 


(b) Following the ideas of Cohen [8], we introduce the functional 

\[ \Psi(z) = A(z^*) - A(z) - \langle A'(z), z^* - z \rangle, \]

where \( z \) is the unique solution of (14) at a certain iteration level and \( z^* \) is the unique solution of the corresponding regularized variational inequality. For example, for \( k = n - 1 \), we have

\[ \Psi(x^n) = A(x_{n-1}) - A(x^n) - \langle A'(x^n), x_{n-1} - x^n \rangle, \]

and for \( k = n \), we have

\[ \Psi(x^{n+1}) = A(x_n) - A(x^{n+1}) - \langle A'(x^{n+1}), x_n - x^{n+1} \rangle. \]

In view of the conditions imposed on the functional \( A \) and on its derivative \( A' \), we have

\[ A(x) - A(z) \geq \langle A'(z), x - z \rangle + \frac{m}{2} \| x - z \|^2 \quad \text{for all} \; x, z \in \mathcal{H}, \tag{20a} \]

\[ A(x) - A(z) \geq \langle A'(x), x - z \rangle - \frac{M}{2} \| x - z \|^2 \quad \text{for all} \; x, z \in \mathcal{H}, \tag{20b} \]

where \( m \) and \( M \) are the moduli of strong monotonicity and Lipschitz continuity of \( A'(-) \), respectively.

To prove that the sequence \( \{x^n\}_{n=1}^{\infty} \) is bounded, we investigate the difference

\[ \Psi(x^n) - \Psi(x^{n+1}) = [A(x_{n-1}) - A(x^n) - \langle A'(x^n), x_{n-1} - x^n \rangle] \]

\[ - [A(x_n) - A(x^{n+1}) - \langle A'(x^{n+1}), x_n - x^{n+1} \rangle] \]

\[ = A(x_{n-1}) - A(x_n) + \langle A'(x^{n+1}), x_n - x_n \rangle \]

\[ + A(x^{n+1}) - A(x^n) - \langle A'(x^n), x_{n-1} - x^n \rangle \]

\[ = A(x_{n-1}) - A(x_n) + \langle A'(x^{n+1}), x_n - x_n \rangle \]

\[ + A(x^{n+1}) - A(x^n) - \langle A'(x^n), x_{n-1} - x^n \rangle. \]

By setting \( x = x^{n+1} \) and \( z = x^n \) in (20a), we obtain

\[ A(x^{n+1}) - A(x^n) - \langle A'(x^n), x^{n+1} - x^n \rangle \geq \frac{m}{2} \| x^{n+1} - x^n \|^2. \]

Moreover, by setting \( x = x_{n-1} \) and \( z = x_n \) in (20b), we obtain

\[ A(x_{n-1}) - A(x_n) \geq \langle A'(x_{n-1}), x_{n-1} - x_n \rangle - \frac{M}{2} \| x_{n-1} - x_n \|^2. \]

In view of the above two inequalities, we have

\[ \Psi(x^n) - \Psi(x^{n+1}) \geq \frac{m}{2} \| x^{n+1} - x^n \|^2 + \langle A'(x_{n-1}), x_{n-1} - x_n \rangle - \frac{M}{2} \| x_{n-1} - x_n \|^2 \]

\[ - \langle A'(x^{n+1}), x_n - x^{n+1} \rangle - \langle A'(x^n), x_{n-1} - x^n \rangle \]

\[ = \frac{m}{2} \| x^{n+1} - x^n \|^2 + \langle A'(x_{n-1}) - A'(x^n), x_{n-1} - x_n \rangle \]

\[ + \langle A'(x^{n+1}) - A'(x^n), x_n - x^{n+1} \rangle - \frac{M}{2} \| x_{n-1} - x_n \|^2. \] \tag{21}

To simplify the notations, we set

\[ T_n(\cdot) = \mathcal{F}(\cdot) + \varepsilon_n(\cdot) \quad \text{and} \quad \tilde{T}_n(\cdot) = \mathcal{F}_n(\cdot) + \varepsilon_n(\cdot). \]

By setting \( z = x_n \) in (19), we obtain

\[ \langle A'(x^{n+1}) - A'(x^n) + \alpha_n \tilde{T}_n(x^n), x_n - x^{n+1} \rangle \geq \alpha_n (\Psi(x^{n+1}) - \Psi(x_n)). \]

Also by setting \( z = x^{n+1} \) in (13), and with \( x_n \) as its unique solution, we obtain

\[ \langle T_n(x_n), x^{n+1} - x_n \rangle \geq \Psi(x_n) - \Psi(x^{n+1}). \] \tag{22}
We combine the above two inequalities to obtain
\[
\langle A'(x^{n+1}) - A'(x^n), x_{e_n} - x^{n+1} \rangle \geq \alpha_n (T_n(x_{e_n}) - \tilde{T}_n(x^n), x_{e_n} - x^{n+1}).
\]

The above inequality, when combined with (21), yields
\[
\Psi(x^n) - \Psi(x^{n+1}) \geq \frac{m}{2} \|x^{n+1} - x^n\|^2 - \frac{M}{2} \|x_{e_n} - x_{e_{n-1}}\|^2
\]
\[+ \alpha_n (T_n(x^n) - T_n(x_{e_n})), x_{e_n} - x_{e_{n-1}}) + \langle A'(x_{e_{n-1}}) - A'(x^n), x_{e_{n-1}} - x_{e_n} \rangle
\]
\[+ \alpha_n (T_n(x^n) - T_n(x_{e_n})), x_{e_n} - x_{e_{n-1}}) := \tau_1 + \tau_2 + \tau_3 + \tau_4 + \tau_5
\]

where the terms \( \tau_i, i = 1, \ldots, 5 \) are defined and estimated below. We have
\[
\tau_1 = \frac{m}{2} \|x^{n+1} - x^n\|^2 + \alpha_n (T_n(x^n) - T_n(x_{e_n})), x_{e_n} - x_{e_{n-1}})
\]
\[\geq \frac{m}{2} \|x^{n+1} - x^n\|^2 - \alpha_n \|T_n(x^n) - T_n(x_{e_n})\| \|x_{e_n} - x_{e_{n-1}}\|
\]
\[\geq \frac{m}{4} \|x^{n+1} - x^n\|^2 - \frac{\alpha_n^2}{m} \|T_n(x^n) - T_n(x_{e_n})\|^2.
\]

In view of the Lipschitz continuity of \( F \), there is a constant \( L \) such that
\[
\|T_n(x^n) - T_n(x_{e_n})\| \leq L \|x^n - x_{e_n}\|.
\]

By plugging this in the above estimate for \( \tau_1 \), we obtain
\[
\tau_1 \geq \frac{m}{4} \|x^{n+1} - x^n\|^2 - \frac{2\alpha_n^2 L^2}{m} \|x^n - x_{e_{n-1}}\|^2 - \frac{2\alpha_n^2 L^2}{m} \|x_{e_n} - x_{e_{n-1}}\|^2.
\]

Furthermore
\[
\tau_2 = \alpha_n (T_n(x^n) - T_n(x_{e_n}), x^n - x_{e_{n-1}})
\]
\[= \alpha_n (T_n(x^n) - T_n(x_{e_{n-1}}), x^n - x_{e_{n-1}}) - \alpha_n (T_n(x_{e_n}) - T_n(x_{e_{n-1}}), x^n - x_{e_{n-1}})
\]
\[\geq \epsilon_n \alpha_n \|x^n - x_{e_{n-1}}\|^2 - \alpha_n \|T_n(x_{e_n}) - T_n(x_{e_{n-1}})\| \|x^n - x_{e_{n-1}}\|
\]
\[\geq \epsilon_n \alpha_n \|x^n - x_{e_{n-1}}\|^2 - \frac{\alpha_n^2}{4} \|x^n - x_{e_{n-1}}\|^2 - L^2 \|x_{e_n} - x_{e_{n-1}}\|^2.
\]

Similarly, we have
\[
\tau_3 = \alpha_n (T_n(x^n) - T_n(x_{e_n}), x_{e_{n-1}} - x_{e_n})
\]
\[= \alpha_n (T_n(x^n) - T_n(x_{e_{n-1}}), x_{e_{n-1}} - x_{e_n}) + \alpha_n (T_n(x_{e_{n-1}}) - T_n(x_{e_n}), x_{e_{n-1}} - x_{e_n})
\]
\[\geq \epsilon_n \alpha_n \|x_{e_n} - x_{e_{n-1}}\|^2 - \alpha_n \|T_n(x^n) - T_n(x_{e_{n-1}})\| \|x_{e_n} - x_{e_{n-1}}\|
\]
\[\geq \epsilon_n \alpha_n \|x_{e_n} - x_{e_{n-1}}\|^2 - \frac{\alpha_n^2}{4} \|x^n - x_{e_{n-1}}\|^2 - L^2 \|x_{e_n} - x_{e_{n-1}}\|^2.
\]

Analogously, we have
\[
\tau_4 = \langle A'(x_{e_{n-1}}) - A'(x^n), x_{e_{n-1}} - x_{e_n} \rangle - \frac{M}{4} \|x_{e_n} - x_{e_{n-1}}\|^2
\]
\[\geq - \|A'(x_{e_{n-1}}) - A'(x^n)\| \|x_{e_n} - x_{e_{n-1}}\|^2 - \frac{M}{4} \|x_{e_n} - x_{e_{n-1}}\|^2
\]
\[\geq - M \|x^n - x_{e_{n-1}}\| \|x_{e_n} - x_{e_{n-1}}\|^2 - \frac{M}{4} \|x_{e_n} - x_{e_{n-1}}\|^2
\]
\[\geq - \frac{\epsilon_n \alpha_n}{2} \|x^n - x_{e_{n-1}}\|^2 - \frac{M^2}{2\epsilon_n \alpha_n} \|x_{e_n} - x_{e_{n-1}}\|^2 - \frac{M}{4} \|x_{e_n} - x_{e_{n-1}}\|^2.
\]
Finally, we have
\[
\tau_5 = \alpha_n (\tilde{T}_n(x^n) - T_n(x^n), x_{n+1} - x_{e_n}) \\
\geq - c\alpha_n \delta_n (1 + \|x^n\|) \|x_{n+1} - x_{e_n}\| \quad \text{(cf. (16))}
\]
\[
\geq - c\alpha_n \delta_n (t + \|x^n - x_{e_{n-1}}\|) \|x_{n+1} - x_{e_n}\| \quad \text{(for any } t \geq (1 + \|x_{e_{n-1}}\|))
\]
\[
\geq - \frac{2\delta^2 \epsilon_n^2}{m} \alpha_n^2 (t + \|x^n - x_{e_{n-1}}\|)^2 - \frac{m \delta_n^2}{8\delta^2} \|x_{n+1} - x_{e_n}\|^2 \quad \text{(for any } \delta \geq \delta_n \forall n)
\]
\[
\geq - s_1 \alpha_n^2 - s_2 \alpha_n^2 \|x^n - x_{e_{n-1}}\|^2 - \frac{m \delta_n^2}{2\delta^2} \|x_{n+1} - x_{e_n}\|^2 - \frac{m}{4} \|x_{n+1} - x^n\|^2 - \frac{m}{2} \|x_{e_{n-1}} - x_{e_n}\|^2
\]
with \(s_1 = \frac{4\epsilon_n^2 \delta_n^2}{m}\) and \(s_2 = \frac{4\epsilon_n \delta_n}{m}\). Here we have used that \(\frac{s_1}{s_2} \leq 1\).

In view of the above estimates, we obtain
\[
\psi(x^n) - \psi(x^{n+1}) \geq \frac{1}{2} \epsilon_n \alpha_n \|x^n - x_{e_{n-1}}\|^2 - c_1 \alpha_n^2 \|x^n - x_{e_{n-1}}\|^2 - c_2 \delta_n^2 \|x^n - x_{e_{n-1}}\|^2
\]
\[
+ \epsilon_n \alpha_n \|x^n - x_{e_{n-1}}\|^2 - c_3 \|x^n - x_{e_{n-1}}\|^2 - c_4 \|x_{n+1} - x_{e_n}\|^2 - s_1 \alpha_n^2.
\]

where \(c_i, i = 1, \ldots, 4\) are positive constants.

Since
\[
c_4 \|x_{n+1} - x_{e_{n-1}}\|^2 \leq \epsilon_n \alpha_n \|x_{n+1} - x_{e_{n-1}}\|^2 + \frac{c_2^2 \|x_{n+1} - x_{e_{n-1}}\|^2}{\epsilon_n \alpha_n}.
\]

After combining the above two inequalities, we obtain
\[
\psi(x^n) - \psi(x^{n+1}) \geq \frac{1}{2} \epsilon_n \alpha_n \|x^n - x_{e_{n-1}}\|^2 - C_1 \kappa_n^2 \|x^n - x_{e_{n-1}}\|^2 - C_2 \|x_{n+1} - x_{e_{n-1}}\|^2 - \epsilon_n \alpha_n \|x_{n+1} - x_{e_n}\|^2
\]
where \(C_1\) and \(C_2\) are positive constants and \(\kappa_n^2 = \alpha_n^2 + \delta_n^2\).

We set
\[
T_n := C_2 \|x_{n+1} - x_{e_{n-1}}\|^2 / \epsilon_n \alpha_n + s_1 \alpha_n^2
\]
and notice that in view of the condition (17), we have \(\sum_{i=1}^{\infty} T_i < \infty\). Finally, we write
\[
\psi(x^{n+1}) - \psi(x^n) \leq C_1 \kappa_n^2 \|x^n - x_{e_{n-1}}\|^2 - \frac{1}{2} \epsilon_n \alpha_n \|x^n - x_{e_{n-1}}\|^2 + T_n
\]
which leads to
\[
\psi(x^{n+1}) \leq \psi(x^n) + \sum_{i=1}^{n} \left\{ C_1 \kappa_i^2 \|x^i - x_{e_{i-1}}\|^2 - \frac{1}{2} \epsilon_i \alpha_i \|x^i - x_{e_{i-1}}\|^2 + T_i \right\}.
\]

By ignoring the term \(-\frac{1}{2} \epsilon_i \alpha_i \|x^i - x_{e_{i-1}}\|^2\) for a while and using (20a), we obtain
\[
\frac{m}{2} \|x^{n+1} - x_{e_n}\|^2 \leq \psi(x^{n+1}) \leq \psi(x^n) + \sum_{i=1}^{n} \left\{ C_1 \kappa_i^2 \|x^i - x_{e_{i-1}}\|^2 + T_i \right\}.
\]

Now, by applying Lemma A.1 (see Appendix), and using the fact that \(\{x_{e_n}\}\) is bounded, we deduce that the sequence \(\{x^n\}\) is bounded.

(c) In view of (19), we have
\[
\langle A(x^{n+1}) - A(x^n) + \alpha_n T(x^n), z - x^{n+1} \rangle \geq \alpha_n (\psi(x^{n+1}) - \psi(z)) \quad \forall z \in K.
\]

By setting \(z = x^n\), we have
\[
\langle A(x^{n+1}) - A(x^n) + \alpha_n T(x^n), x^n - x^{n+1} \rangle \geq \alpha_n (\psi(x^{n+1}) - \psi(x^n)).
\]
The above inequality, in view of the boundedness of \( \{x^n\} \), implies that
\[
m\|x^{n+1} - x^n\|^2 \leq \langle A'(x^{n+1}) - A'(x^n), x^{n+1} - x^n \rangle
\leq \alpha_n \langle T(x^n), x^n - x^{n+1} \rangle + \alpha_n \{\psi(x^n) - \psi(x^{n+1})\}
\leq \alpha_n (\|\tilde{T}(x^n)\| + \ell)\|x^{n+1} - x^n\|.
\]

The assertion now follows from the above inequality.

(d) Now we take the functional \( A \) to be as in (18). In this case, we have
\[
\langle A'(x), y \rangle = a(x, y) \quad \text{for all } x, y \in \mathcal{H}.
\]
Let us introduce an inner-product on \( \mathcal{H} \) as follows:
\[
\langle x, y \rangle_N = a(x, y) \quad \text{for all } x, y \in \mathcal{H}.
\]
Let \( \cdot \| \cdot \| \) be the norm induced by \( \langle x, y \rangle_N \). Then we have
\[
a_1 \|x\| \leq \|x\|_N \leq b_1 \|x\| \quad \text{for all } x, y \in \mathcal{H}
\]
where \( a_1 \) and \( b_1 \) are the moduli of coercivity and continuity for \( a(\cdot, \cdot) \). Therefore, the norm \( \| \cdot \|_N \) is equivalent to the original norm \( \| \cdot \| \) of \( \mathcal{H} \).

By setting \( z = x_{\epsilon_n} \) in (19), we obtain
\[
\langle A'(x^{n+1}) - A'(x^n) + \alpha_n \tilde{T}(x^n), x_{\epsilon_n} - x^{n+1} \rangle \geq \alpha_n (\psi(x^{n+1}) - \psi(x_{\epsilon_n})).
\]

The above inequality, when combined with (22), yields
\[
\|x^{n+1} - x_{\epsilon_n}\|_N^2 = \langle x^{n+1} - x_{\epsilon_n}, x^{n+1} - x_{\epsilon_n} \rangle_N
= \langle A'(x^{n+1}) - A'(x_{\epsilon_n}), x^{n+1} - x_{\epsilon_n} \rangle
\leq \langle A'(x_{\epsilon_n}) - A'(x^n), x_{\epsilon_n} - x^{n+1} \rangle + T_1,
\]
where \( T_1 = \alpha_n \langle \tilde{T}(x^n) - T_n(x_{\epsilon_n}), x_{\epsilon_n} - x^{n+1} \rangle \).

This further implies that
\[
\|x^{n+1} - x_{\epsilon_n}\|_N^2 \leq \|x^n - x_{\epsilon_n}\|_N^2 + 2T_1
= \|x^n - x_{\epsilon_n-1}\|_N^2 + \|x_{\epsilon_n-1} - x_{\epsilon_n}\|_N^2 - 2\langle x^n - x_{\epsilon_n-1}, x_{\epsilon_n} - x_{\epsilon_n-1} \rangle_N + 2T_1
\leq \|x^n - x_{\epsilon_n-1}\|_N^2 + C_1 \|x_{\epsilon_n-1} - x_{\epsilon_n}\|_N + 2T_1
\]
where \( C_1 \) is a constant. For an upper bound for \( T_1 \), we have
\[
2T_1 = 2\alpha_n \langle \tilde{T}(x^n) - T_n(x^n), x_{\epsilon_n} - x^{n+1} \rangle + 2\alpha_n \langle T_n(x^n) - T_n(x_{\epsilon_n}), x_{\epsilon_n} - x^n \rangle
\leq 2c_n \alpha_n \delta_0 (1 + \|x^n\|) \|x_{\epsilon_n} - x^n\| + 2\alpha_n \langle T_n(x^n) - T_n(x_{\epsilon_n}), x_{\epsilon_n} - x^n \rangle
\leq -2\epsilon_n \alpha_n b_1^{-2} \|x^n - x_{\epsilon_n}\|_N^2 + 2C_1 \epsilon_n \alpha_n b_1^{-2} \|x_{\epsilon_n-1} - x_{\epsilon_n}\|_N + 2C_2 \alpha_n \delta_0 + 2\alpha_n \delta_n
\]
where \( C_2 \) is a constant.

We combine the above two estimates and obtain
\[
\|x^{n+1} - x_{\epsilon_n}\|_N^2 \leq \|x^n - x_{\epsilon_n-1}\|_N^2 - 2\epsilon_n \alpha_n b_1^{-2} \|x^n - x_{\epsilon_n-1}\|_N^2 + A_n
\]
where \( A_n := C_3 (\|x_{\epsilon_n-1} - x_{\epsilon_n}\|_N + \alpha_n^2 + \alpha_n \delta_n) \) and \( C_3 \) is a suitable constant.

In view of the imposed conditions, we have \( \frac{A_n}{\epsilon_n x_{\epsilon_n}} \to 0 \) and this, in view of Lemma A.2 (see Appendix), implies that \( \|x^{n+1} - x_{\epsilon_n}\|_N \to 0 \). The assertion now follows from the equivalence of \( \| \cdot \|_N \) and \( \| \cdot \| \), and the fact that \( x_{\epsilon_n} \to x^* \). \( \square \)
4. A numerical example

In a forthcoming work, we will present detailed numerical experiments and a comparison between iterative regularization and proximal regularization. For now, we will just demonstrate the feasibility of the approach by using (7) for estimating a coefficient in the scalar BVP (1).

We take $\Omega = [0, 1] \times [0, 1]$. The exact coefficient we try to identify is $a = 1 + xy^2$. The exact solution is $u = xy(1-x)(1-y)$ and the source term is $f = -y^3 + y^4 + 4y^3x - 4y^4x + 2y - 2y^2 - 2x^2y + 6x^2y^2 + 2x^3y - 6x^3y^2 + 2x - 2x^2$.

We used finite element discretization, and in all the finite-element computations for this example, we used a sequence of uniform triangulations on $\Omega$. We employed piecewise quadratic elements to represent the solution and piecewise linear elements to represent the coefficients. We solved the minimization problem (7) for the modified output least-squares functional by using $A(x) = \|x\|^2$ in (14), that is, by the projected gradient analogue (15). (In fact, due to the slow speed, we took the liberty of using a scaled projected gradient algorithm instead.) We remark that since this a synthetic experiment, the data vector is computed (by numerically solving the direct problem); it is not measured. The results are shown in Fig. 1.

5. Concluding remarks

We studied an abstract elliptic inverse problem by using the modified output least-squares functional. Since this functional is convex, the regularized problem leads to a strongly convex objective functional. A variational inequality formulation was used to suggest and analyze an algorithm for its numerical solution. The convergence analysis depends on the convexity of the functional, and the main result is proved under much weaker conditions than the classical output least-squares functional would require. A numerical example is given that shows the feasibility of the approach.
Acknowledgements

We are grateful to the referees for a careful reading and suggestions that brought a substantial improvement. A part of this work was carried out during the first author’s stay at the Department of Mathematics, Technical University of Kaiserslautern, Kaiserslautern, Germany, during the period of August–September 2006. This author is greatly thankful to Dr. Falk Triebsch for providing an excellent working environment and to DAAD for providing financial assistance. Special thanks are due to Ms. Sheila Berry, whose vital help made this trip to Germany possible.

Appendix

We recall the following results.

Lemma A.1. Let \( \{r^k\}_{k \in \mathbb{N}} \) and \( \{s^k\}_{k \in \mathbb{N}} \) be two sequences of positive real numbers such that \( \sum_{k \in \mathbb{N}} s^k < +\infty \) and \( r^{k+1} \leq \sum_{l=1}^{k} s^l r^l + t^k \), where \( t^k \leq t \), for all \( k \in \mathbb{N} \). Then the sequence \( r^k \) is bounded.

Proof. See Lemma 5 in [9]. \( \square \)

Lemma A.2 ([2, Lemma 7.1.2]). Let \( \{a_n\} \) and \( \{c_n\} \) be sequences of non-negative real numbers, \( \{b_n\} \) be a sequence of positive real numbers satisfying the inequality

\[
a_{n+1} \leq (1 - b_n)a_n + c_n, \quad b_n \leq 1
\]

where \( b_n \to 0 \) as \( n \to 0 \), \( \sum_{n=1}^{\infty} b_n = \infty \), and \( \lim \frac{c_n}{b_n} = 0 \). Then the sequence \( \{a_n\} \) converges to 0.

References