Gradient-type methods in inverse parabolic problems

Sergey Kabanikhin and Aleksey Penenko

Abstract. This article is devoted to gradient-based methods for inverse parabolic problems. In the first part, we present a priori convergence theorems based on the conditional stability estimates for linear inverse problems. These theorems are applied to backwards parabolic problem and sideways parabolic problem. The convergence conditions obtained coincide with sourcewise representability in the self-adjoint backwards parabolic case but they differ in the sideways case. In the second part, a variational approach is formulated for a coefficient identification problem. Using adjoint equations, a formal gradient of an objective functional is constructed. A numerical test illustrates the performance of conjugate gradient algorithm with the formal gradient.

1. Introduction
We investigate several inverse and ill-posed problems for heat conduction equations which can be reduced to the operator equation

$$Aq = f.$$ (1)

The idea of replacing (1) by the variation problem

$$J(q) = \|Aq - f\|^2 \rightarrow \min$$ (2)

dates back to A.M. Legendre (1806), K. Gauss (1809) and L.V. Kantorovich (1948).

Here we are concerned with the analysis of gradient-based algorithms applied to the minimization problem stated. Namely, in section 2, we present convergence theorems based on the conditional stability estimates for linear inverse parabolic problems. In section 3, a formal gradient for the nonlinear inverse problem is constructed. The usage of this formal gradient in a conjugate gradient method is illustrated by a numerical experiment.

2. Linear inverse heat conduction problems
We consider the case where $A : Q \rightarrow F$ is a linear operator acting between two separable Hilbert spaces and therefore $J(q)$ is differentiable and $J'(q) = 2A^*(Aq - f)$.

Lemma. Assume that $A$ is linear and bounded and the solution of (1) is unique. Then $J(q)$ has no more than one stationary point.

In nonlinear case Lemma holds true if the condition $J'(q) = 0$ implies that $Aq = f$.

We consider several gradient-type methods,

$$q_{n+1} = q_n - \alpha J'q_n$$ (3)

including Landweber iteration ($\alpha \in (0, \|A\|^{-2})$), steepest descent and others and estimate the convergence rate using the conditional stability function [1–2],
2.1. Strong convergence based on the conditional stability estimates

**Definition.** The decreasing real-valued function $\beta(\delta)$ is said to be a conditional stability function for the operator $A$ on a set $M \subset Q$ if the following two conditions are valid:
\[
\lim_{\delta \to +0} \beta(\delta) = 0 \quad \text{and} \quad \sup_{f_1, f_2 \in A(M)} \|f_1 - f_2\| \leq \delta \quad \text{the inequality} \quad \|A^{-1}f_1 - A^{-1}f_2\| \leq \beta(\delta)
\]

**Theorem** (the weak convergence rate). Let a solution $q_e$ to the equation $Aq = f$ exist. Then the sequence defined by (3) satisfies the inequality
\[
J(q_n) \leq \frac{\|q_e - q_0\|^2}{\alpha n(1 - \alpha\|A\|^2)}.
\]

**Theorem** (the strong convergence rate). Let $f \in A(M)$ and $q_e$ be a solution of $Aq = f$. Here $M = \{q \in Q : \|q\| \leq 2\|q_e\|\}$. Assume that $\beta(\delta)$ is a conditional stability function for the operator $A$ on the set $M$ and $q_0 \in M$. Then the sequence (3) satisfies the inequality
\[
\|q_n - q_e\| \leq \beta\left(\frac{\|q_0 - q_e\|}{\sqrt{\alpha n(1 - \alpha\|A\|^2)}}\right).
\]

Suppose that instead of the exact data $f \in A(M)$ we have its approximation $f_\delta \in F$ which satisfies the condition
\[
\|f - f_\delta\| \leq \delta.
\]

The case when $f_\delta$ does not belong to $A(M)$ is the most complicated. We introduce the corresponding misfit function
\[
J_\delta(q) = \|Aq - f_\delta\|^2
\]
and use the method of Landweber iteration:
\[
q_{\delta,n+1} = q_{\delta,n} - \alpha J'q_{\delta,n}, \quad \alpha \in (0, \|A\|^{-2}), \quad q_{\delta,0} = q_0.
\]

Notice that the sequences (3) and (7) have the same initial guess $q_0$ and the same parameter $\alpha$. We prove that
\[
\|q_n - q_{\delta,n}\| \leq \beta_1(n)\delta,
\]
where
\[
\beta_1(n) = 2\alpha\|A\|\sum_{k=1}^{n} \|I - 2\alpha A^*A\|^{k-1} = 2\alpha\|A\|\frac{1 - \|I - 2\alpha A^*A\|^{n+1}}{1 - \|I - 2\alpha A^*A\|}.
\]

**Theorem** Let $A : Q \to F$ be a linear continuous operator. Suppose that for some $f \in F$, there exists $q_e$ solving equation (1). Let $\beta(\delta)$ be a conditional stability function for the operator $A$ on the set $M = \{q \in Q : \|q\| \leq 2\|q_e\|\}$.

We assume that $f_\delta \in F$ satisfies (5), $q_0 \in M$ and a sequence $\{q_n\}$ is defined by (7). Then for all integer $n$ the following estimate holds true:
\[
\|q_n - q_{\delta,n}\| \leq \beta\left(\frac{\|q_e - q_0\|}{\sqrt{\alpha n(1 - \alpha\|A\|^2)}}\right) + \delta \beta_1(n).
\]

2.2. Backwards parabolic equation

Let $D \subset \mathbb{R}^m$ be a connected bounded domain with a smooth boundary and $L(y)$ is the elliptic self-adjoint operator of the second order.

Then if a solution to the problem
\[
u_t = -L(y)u, \quad y \in D, \quad t \in (0, T),
\]
exists and is smooth enough, it can be shown that, for all $t \in (0, T)$,
\[
\int_D u^2(y, t) \, dy \leq \|q\|_{L_2(D)}^{2T/T} \|f\|_{L_2(D)}^{2(T-t)/T},
\]
where $q(y) = u(y, T)$.

We define the discrepancy function as
\[
J(q) = \int_D [u(y, 0) - f(y)]^2 \, dy.
\]

Using $q_n$, we can find the solution to the direct problem
\[
u^n_t = -L(y)u^n, \quad y \in D, \quad t \in (0, T),
\]
\[
u^n|_{\partial D} = 0, \quad q^n(y) = u^n(y, T),
\]
and then estimate the difference $u - u^n$ according to (12)
\[
\int_D (u - u^n)^2(y, t) \, dy \leq \|q - q_n\|_{L_2(D)}^{2T/T} \|f - Aq_n\|_{L_2(D)}^{2(T-t)/T}.
\]

We cannot estimate the rate of convergence of $\|q - q_n\|_{L_2(D)}$. But we do estimate the rate of the strong convergence of $\int_D (u - u^n)^2(y, t) \, dy$ for every $t \in (0, T)$. Notice that the less $t$ the stronger the mentioned rate of convergence.

2.3. Sideways heat conduction equation
Consider the problem
\[
u_t = \nu_{xx}, \quad x \in [0, 1], \quad t \in (0, \infty),
\]
\[
u(x, 0) = 0, \quad x \in [0, 1],
\]
\[
u(0, t) = q(t), \quad t \in [0, \infty],
\]
\[
u_x(1, t) = 0, \quad t \in [0, \infty],
\]
\[
u(1, t) = f(t), \quad t \in [0, \infty],
\]
with unknown function $q(t)$.

In this case, the following conditional stability estimate holds [4] for $q(t), f(t) \in L_2([0, \infty])$:
\[
\|\nu(x, \cdot)\|_{L_2([0, \infty])} \leq C_1\|f\|_{L_2([0, \infty])} + C_2\|q\|_{L_2([0, \infty])}^{1-x} \|f\|_{L_2([0, \infty])}^x, \quad x \in [0, 1].
\]

The discrepancy function is defined in the same way as in the previous subsection:
\[
J(q) = \int_0^\infty [\nu^n(1, t) - f(t)]^2 \, dt,
\]
where $u^n(x, t)$ is the solution of the direct problem
\[
u^n_t = \nu^n_{xx}, \quad x \in [0, 1], \quad t \in (0, \infty),
\]
\[
u^n(x, 0) = 0, \quad x \in [0, 1],
\]
\[
u^n(0, t) = q^n(t), \quad t \in [0, \infty],
\]
Let \( q_n \) be a minimizing sequence for the functional \( J(q) \). In this case, according to conditional stability estimate

\[
\int_0^\infty [u(x, t) - u^n(x, t)]^2 dt \leq C_1 \sqrt{J(q_n)} + C_2 \|q - q_n\|_{L_2([0, \infty])}^{1-x} \left[ \sqrt{J(q_n)} \right]^x, \quad x \in [0, 1],
\]

one obtains convergence in the inner points of the spatial domain \([0, 1]\).

We should note the difference of the a priori convergence condition provided by our approach from the sourcewise representability condition (see e.g. \([3]\)).

Sourcewise representability condition has the form

\[
q = (A^*A)^\top w, \quad \|w\| \leq \rho,
\]

and the difference will be evident in comparison of both the backwards parabolic problem and the sideways heat conduction problem. For the backwards parabolic problem our approach guarantees strong convergence of \( \int_{0}^{T} (u - u^n)^2(y, t)dy \) for every \( t \in (0, T) \). The condition that \( t \) is inside the interval \((0, T)\) basically means for the selfadjoint operator \( A \) that

\[
u(y, t) - u^n(y, t) = A^\top(q - q_n) = (A^*A)^{1/2} (q(y) - q_n(y)).
\]

This is structurally similar to the sourcewise representability condition.

On the contrary, in the sideways heat conduction equation, sourcewise representability condition still has the form

\[
q(t) = (A^*A)^\top w, \quad \|w\| \leq \rho,
\]

while according to our approach it is enough for \( x \) in \( \int_0^\infty [u(x, t) - u^n(x, t)]^2 dt \) to be just inside the spatial interval \([0, 1]\) which seems to be more physically feasible than sourcewise representability.

3. Coefficient identification problem

The theory presented in the previous section is valid for linear inverse problems for parabolic equations. In general, the variational approach can also be applied to the nonlinear inverse parabolic problems namely to coefficient identification problem for heat conduction equation.

3.1. Problem statement

Let us consider the boundary value problem in the domain \( \Omega_T = [0, 1] \times [0, T] \):

\[
u_t(x, t) - (k(x)u_x(x, t))_x = 0, \quad t \in [0, T], x \in [0, 1],
\]

\[
k(0)u_x(0, t) = \alpha_L(t), \quad t \in [0, T],
\]

\[
k(1)u_x(1, t) = \alpha_R(t), \quad t \in [0, T],
\]

\[
u(x, 0) = \varphi(x), \quad x \in [0, 1].
\]

The problem of solving (26)-(29) for the function \( u(x, t; k) \) with known functions \( \alpha_L(t), \alpha_R(t), \varphi(x), k(x) \) will be referred to as the direct problem.

The problem of finding \( k(x) \), such that \( u(x, t; k) \) solves (26)-(29) and satisfies conditions:

\[
u(0, t; k) = f_L(t), \quad t \in [0, T],
\]

\[
u(1, t; k) = f_R(t), \quad t \in [0, T],
\]

will be referred to as the inverse one. Functions \( \alpha_L(t), \alpha_R(t), \varphi(x), f_L(t), f_R(t) \) are considered to be known.
3.2. The formal gradient
In this problem we also apply variational technique to obtain least squares norm solution of the inverse problem. To do this, we introduce a discrepancy function

$$J(k) = \beta_L \int_0^T (u(0, \tau; k) - f_L(\tau))^2 d\tau + \beta_R \int_0^T (u(1, \tau; k) - f_R(\tau))^2 d\tau.$$  (32)

This function is defined on

$$K = \{k(x) \in L_\infty(0, 1) : 0 < k_{min} \leq k(x) \leq k_{max}\} \subset L_2(0, 1).$$

Minimization of this function is implemented with conjugate gradient method. To do this, one has to derive the gradient first. It can be accomplished with the help of the adjoint problem approach. This approach had been applied to the more general meteorological models in e.g. [5] and to this specific problem in [6] - [10]. The following theorem describes the gradient (as in [10]) and the residual term:

**Theorem [10], [11]** Let $k(x), k(x) + \delta k(x) \in K$, $\varphi(x) \in H^1(0, 1)$, $\alpha_L(t) \in H^1(0, T)$, $\alpha_R(t) \in H^1(0, T)$. $u(x, t; k)$ be the weak solution of the direct problem (26)-(29) and $\delta u(x, t; k, \delta k) = u(x, t; k + \delta k) - u(x, t; k)$ then

$$J(k + \delta k) - J(k) = \left\langle - \int_0^T u_x(., \tau; k) \psi_x(., \tau; k) d\tau, \delta k(.) \right\rangle_{L_2(\Omega_T)} + O(\|\delta k(.)\|_{L_\infty(\Omega_T)}^2).$$  (33)

Here $\psi(x, t; k)$ is the weak solution in $V^{1,0}(\Omega_T)$ of the adjoint problem

$$\psi_t(x, t) + (k(x) \psi_x(x, t))_x = 0,$$  (34)
$$k(0) \psi_x(0, t) = -2 \beta_L (u(0, t; k) - f_L(t)),$$  (35)
$$k(1) \psi_x(1, t) = 2 \beta_R (u(1, t; k) - f_R(t)),$$  (36)
$$\psi(x, T) = 0.$$  (37)

Rigorously speaking, the gradient obtained is a formal one because the Cartesian product in (33) is from $L_2((0, 1))$ while the residual is of the second order of magnitude with respect to the $\delta k(.)$ norm in $L_\infty((0, 1))$. This formal gradient is denoted by

$$\tilde{J}(k)(x) = - \int_0^T u_x(x, \tau; k) \psi_x(x, \tau; k) d\tau.$$  (38)

3.3. The conjugate gradient method
The subsequent steps are quite common for the conjugate gradient minimization procedures.

Suppose that the approximation $k^{[l]}$ to the solution is already obtained.

(i) $J(k^{[l]})$ is calculated.
(ii) If $J(k^{[l]})$ or $l$ does not fulfill the stopping conditions of the iterations, then the formal gradient $\tilde{J}(k^{[l]})(x)$ is calculated.
(iii) The descent direction is chosen according to the rule

$$p_l(x) = \begin{cases} \tilde{J}(k^{[l]})(x), l = 1 \\ \tilde{J}(k^{[l]})(x) + \frac{\|\tilde{J}(k^{[l-1]})(x)\|^2}{\|\tilde{J}(k^{[l-1]})(x)\|^2} p_{l-1}(x), l > 1 \end{cases}$$
(iv) The descent parameter $d^{[l]}$ is chosen.
(v) The next approximation $k^{[l+1]}(x) = k^{[l]}(x) - d^{[l]}p_l(x)$ is calculated.

The only peculiarity is the choice of the descent parameter $d^{[l]}$. The formula that is used for the case of linear operators fails here. That is why, we use the golden ratio search to choose the descent parameter for the conjugate gradient iterations. This is the improved modification of the dichotomy method for solving minimization problems with a continuous functional

$$J(k^{[l]}(x) - d^{[l]}p_l(x)) \rightarrow \min$$

with respect to $d^{[l]}$.

Figures 1,2 and 3 present a result of a numerical experiment with the obtained formal gradient. In this experiment $\alpha_L(t) = \delta(t), \alpha_R(t) = 0, \varphi(x) = 0$ and $f_R(t)$ is unknown. The latter fact resulted in the worsened reconstruction on the right boundary (see fig.3).

**Figure 1.** Relative error in $L_\infty(0,1)$ of the coefficient identification problem with the respect to the iteration number.

**Figure 2.** $\log(J(k^{[l]}))$ of the coefficient identification problem with respect to the iteration number.

**Figure 3.** Exact and reconstructed solutions for the coefficient identification problem

**Acknowledgments**

The work was supported by the SB RAS (Integration grants 16 and 48) and by Russian Foundation for Basic Research grants No 08-01-00312-a and No 08-01-90260-Uzb.

**References**

[1] Kabanikhin S I 2005 Convergence rate estimation of gradients methods via conditional stability of inverse and ill-posed problems *J. Inv. Ill-Posed Problems* **13** No. 3 259–64
[2] Kabanikhin S I 2006 Conditional stability stopping rule for gradient methods applied to inverse and ill-posed problems J. Inv. Ill-Posed Problems 14 No. 8 805–12

[3] Engl H, Hanke M and Neubauer A 1996 Regularization of Inverse Problems (Netherlands: Kluwer Academic Publishers)

[4] Knabner P and Vesella S 1988 The optimal stability estimate for some ill-posed Cauchy problems for a parabolic equation. Math. Methods in the Applied Sciences 10 575–83

[5] Penenko V V 1975 Computational issues of Atmospheric processes dynamics and estimates of various factors impact on the atmosphere dynamics (Some problems of numerical and applied mathematics) ed M.M. Lavrentiev (Novosibirsk,"Nauka") pp 61–77

[6] DuChateau P 1995 Monotonicity and invertibility of coefficient-to-data mappings for parabolic inverse problems SIAM J. Math. Anal. 26 1473–87

[7] DuChateau P 1996 Introduction to inverse problems in partial differential Equations for engineers, Physicists and Mathematicians (Parameter Identification and Inverse Problems in Hydrology, Geology and Ecology) eds J Gottlieb and P DuChateau (Netherlands: Kluwer Academic Publishers) pp 3–38

[8] DuChateau P 1997 An inverse problem for the hydraulic properties of porous media, SIAM J. Math. Anal. 28 611–32

[9] DuChateau P, Thelwell R and Butters G 2004 Analysis of an adjoint problem approach to the identification of an unknown diffusion coefficient Inverse Problems 20 601–25

[10] Hasanov A, DuChateau P and Pektas B 2006 An adjoint problem approach and coarse-fine mesh method for identification of the diffusion coefficient in a linear parabolic equation J. Inv. Ill-Posed Problems 14 No. 4 1–29

[11] Kabanikhin S I, Hasanov A and Penenko A V 2008 The gradient-based method for solving the inverse coefficient heat-conduction problem Siberian J. Num. Math. / Sib. Branch of Russ. Acad. of Sci. 11 No. 1 41–54