REGULARIZING GRADIENT ALGORITHMS FOR INVERSE THERMAL-CONDUCTION PROBLEMS

O. M. Alifanov and S. V. Rumyantsev

UDC 536.24

Regularizing algorithms are presented for solving linear inverse problems by steepest-descent, minimum-discrepancy, and conjugate-gradient methods.

An automated system for processing data from heat experiments imposes stringent requirements on the processing algorithms; the most important of these are reasonable universality and independence in the algorithms, which have to handle a range of jobs in the processing of measurement data and other information without intervention from the operator. These specifications are extremely important in the case of inverse boundary-value problems in thermal conduction, on account of the irregularity of the formulation. Therefore, algorithms free from the deficiencies of the subjective approach in the derivation of the functions are required that are suitable for automatic systems, and they should be constructed via the theory of methods of solving inverse problems [1]. From the mathematical viewpoint, a boundary-value problem in thermal conduction involves an equation of the first kind for bodies whose thermophysical parameters are independent of temperature:

$$Au = \bar{f}, \quad u \in U, \quad \bar{f} \in F, \tag{1}$$

where A: U → F is a linear continuous operator; U, F are Hilbert spaces; and the inverse operator A^{-1} is unbounded. We assume that the solution to (1) is not in general unique.

If (1) had a bounded inverse operator, one could use gradient minimization methods in the solution; such methods are attractive in that they enable one to handle unconditional minimization as well as minimization subject to constraints, which is particularly important for inverse problems, where proper solution is dependent on proper use of a priori information.

However, gradient methods cannot be applied directly to incorrectly formulated problems because errors on the right sides and in the operator in (1) may result in divergent successive approximations, although it has been shown [2-5] that the corresponding sequence converges to the normal solution of (1) on exact data. Therefore, it is necessary to examine whether one can modify gradient methods in such a way as to obtain stable algorithms. It has been suggested [6] that a regularizing algorithm can be based on simple iteration, in which the regularization consists in choosing the number of iterations to suit the error of the input data. A difference from a correctly formulated problem is that the number of iterations cannot be arbitrarily large, but instead must be restricted to a certain number, which is less the greater the error in the input data. This is the only approach to the construction of regularizing algorithms via iterative methods and has been applied [7, 8] to boundary-value problems in thermal conduction, although no strict mathematical demonstration was given. We have provided a basis for the approach and have derived the conditions for matching the number of iterations to the error of the input data in the various gradient methods.

We assume that the operator A and right side $\bar{f}$ of (1) are known approximately, i.e., we are given a linear operator A_h from some family of operators that approximate to A subject to the approximation condition $\|A - A_h\| \leqslant h$, together with the right side $f_\delta = \bar{f} + \tilde{f}$, where $\tilde{f} \in F$ is some noise process; $\|\tilde{f}\|_F = \delta$. By $U_{\bar{f}}$ we denote the set of solutions to (1) on exact data. The following theorem indicates what properties should occur in an iterative algorithm in order that it may be regularized by matching the number of iterations to the error $\sigma = \{\delta, h\}$ of the input data.

THEOREM 1. Let $u_{n+1} = F(u_n, \tilde{f}, h)$ be some iterative algorithm having the following features:

1) $\forall u_0 \in U$ for the sequence $\{\bar{u}_n\}$ derived on exact data (i.e., $\bar{u}_{n+1} = F(\bar{u}_n, 0, 0)$) converges to some element in a norm of space U and $F(u, 0, 0) = u$, $\forall u \in U_{\bar{f}}$;

2) the mapping F(u, f̃, h) is continuous at all points $\{u, 0, 0\}$, apart perhaps from the points $\{\tilde{u}, 0, 0\}$, $\tilde{u} \in U_{\bar{f}}$.

Then $\exists \delta_0 > 0$, $h_0 > 0 : \forall \sigma = \{\delta, h\} : \|\tilde{f}\|_F = \delta \leqslant \delta_0$, $h \leqslant h_0$, $\exists N(\sigma) : \|u_{N(\sigma)} - \bar{u}\|_U \underset{\sigma \to 0}{\to} 0$.

The proof of the theorem is based on considering two sequences of approximations starting from the same estimate u_0: $\{u_{n+1} = F(u_n, \tilde{f}, h)\}$ a sequence on perturbed data and $\{\bar{u}_n\}$ a sequence on exact data. Then we estimate the deviation of

Translated from Inzhenerno-Fizicheskii Zhurnal, Vol. 39, No. 2, pp. 253-258, August, 1980. Original article submitted January 2, 1980.

0022-0841/80/3902-0858$07.50

the elements u_n from $\bar{u}_n$ by means of 2) for algorithm F and use the convergence of sequence $\{\bar{u}_n\}$ (property 1).

Theorem 1 can be applied to examine the regularizability of any iterative algorithm; as examples we consider the following gradient methods:

$$u^0_{n+1} = u^0_n - \beta^0_n J'u^0_n, \quad \beta^0_n = \frac{\Delta^2_n}{\|J'u^0_n\|^2_U}, \tag{2}$$

which is steepest descent, and

$$u^m_{n+1} = u^m_n - \beta^m_n J'u^m_n, \quad \beta^m_n = \frac{\|J'u^m_n\|^2_U}{\|A_h J'u^m_n\|^2_F}, \tag{3}$$

which is minimal discrepancy, as well as

$$u^C_{n+1} = u^C_n - \beta^C_n p_n, \quad \beta^C_n = \frac{(J'u^C_n, p_n)_U}{\|A_h p_n\|^2_F}, \quad p_n = J'u^C_n + \gamma_{n-1}p_{n-1},$$
$$p_0 = J'u_0, \quad \gamma_{n-1} = \frac{\|J'u^C_n\|^2_U}{\|J'u^C_{n-1}\|^2_U}, \tag{4}$$

which is the conjugate-gradient method. Here $J'u = A^*_h(A_h u - f_\delta)$ is the gradient in the square of the discrepancies and $\Delta_n = \|A_h u_n - f_\delta\|_F$ is the discrepancy at iteration n.

Methods (2), (3), and (4) satisfy condition 1) of Theorem 1 [2-5]. It can be shown that condition 2) of Theorem 1 will also be obeyed for each of these methods. For this purpose it is necessary to establish that the gradient $J'(u, \tilde{f}, h)$ and the discrepancy $\Delta(u, \tilde{f}, h)$ are continuous at all points $\{u, 0, 0\}$. Also, $J'(u, 0, 0) = 0$ and $\Delta(u, 0, 0) = 0, \Longleftrightarrow u \in U_{\bar{f}}$. From the orthogonality of the transform of the conjugate operator $R(A^*) = \{u \in U : u = A^*f, f \in F\}$ for the core of operator A $N(A) = \{u \in U : Au = 0\}$ it follows that $AJ'(u, 0, 0) = 0 \Longleftrightarrow J'(u, 0, 0) = 0$. These arguments allow us to prove that condition 2) is met for methods (2)-(4). Then the following applies:

THEOREM 2. For each of the methods of (2)-(4) there exist ways of choosing the number $N(\sigma)$ of iterations that will regularize methods (2)-(4).

We now consider detailed ways of choosing the number of iterations.

THEOREM 3. Methods (2), (3), and (4) are regularizing algorithms if the numbers of iterations are chosen in accordance with the conditions

$$N_0(\sigma) = \min\{n : \Delta_{n+1} < 2\Delta_e = 2(\delta + h\|\bar{u}\|_U)\}, \tag{5}$$

$$N_m(\sigma) = \min\{n : \frac{\Delta^2_{n+1} + \Delta^2_{n+2}}{2\Delta_{n+1}} < \Delta_e\}, \tag{6}$$

$$N_c(\sigma) = \min\left\{n : \frac{(\Delta^2_{n+1} + \Delta^2_{n+2})\sum_{i=0}^{n+1} \frac{1}{\|J'u^C_i\|^2_U}}{2\sum_{i=0}^{n+1} \frac{\Delta^2_i}{\|J'u^C_i\|_U}} < \Delta_e\right\} \tag{7}$$

The proof of Theorem 3 is based on Theorem 2 and on the following auxiliary assertions.

LEMMA 1. $J'(u, \tilde{f}, h) \neq 0$ for all $u \in D_e$, where $D_e = \{u \in U : \|A_h u - f_\delta\| \leqslant \min_{\bar{u} \in U_{\bar{f}}}(\delta + h\|\bar{u}\|_U)\}$.

LEMMA 2. The following inequalities apply respectively for methods (2), (3), and (4):

$$\|u^0_n - \bar{u}\|^2_U - \|u^0_{n+1} - \bar{u}\|^2_U = \beta^0_n(\Delta^2_n + 2((A - A_h)\bar{u} + \tilde{f}, A_h u_n - f_\delta)_F) \geqslant \beta^0_n \Delta_n(\Delta_n - 2\Delta_e), \tag{8}$$

$$\|u^m_n - \bar{u}\|^2_U - \|u^m_{n+1} - \bar{u}\|^2_U = \beta^m_n(\Delta^2_n + \Delta^2_{n+1} + 2((A - A_h)\bar{u} + \tilde{f},$$
$$A_h u_n - f_\delta)_F) \geqslant \beta^m_n \Delta_n\left(\frac{\Delta^2_n + \Delta^2_{n+1}}{\Delta_n} - 2\Delta_e\right), \tag{9}$$

$$\|u_n^C-\bar{u}\|_U^2-\|u_{n+1}^C-\bar{u}\|_U^2=\beta_n^C\Big((\Delta_n^2+\Delta_{n+1}^2)\sum_{i=0}^{n}\frac{\|J'u_n^C\|_U^2}{\|J'u_i^C\|_U^2}+$$

$$+2\sum_{i=0}^{n}\frac{\|J'u_n^C\|_U^2((A-A_h)\bar{u}+\tilde{f},\,A_hu_i^C-f_\delta)_F}{\|J'u_i^C\|_U^2}\Big)\geqslant\left(\Delta_n^2+\Delta_{n+1}^2-2\Delta_e\frac{\sum_{i=0}^{n}\frac{\Delta_i}{\|J'u_i^C\|_U^2}}{\sum_{i=0}^{n}\frac{1}{\|J'u_i^C\|_U^2}}\right)\beta_n^C\sum_{i=0}^{n}\frac{\|J'u_n^C\|_U^2}{\|J'u_i^C\|_U^2}. \qquad (10)$$

Inequalities (8)-(10) explain the meanings of the methods of halting the iteration presented in Theorem 3: in any iteration whose number does not exceed the number defined by (5)-(7), one gets approximations that tend monotonically to the solution to (1). One cannot guarantee that better approximations will be obtained with larger numbers of iterations. An exception is constituted by the conjugate-gradient method, where condition (7) can be used not to halt the process but to re-initiate it. In fact, it is readily shown that $\Delta_0\geqslant\Delta_{N_C}>\Delta_e$ (this follows from the fact that Δ_n decreases monotonically). Then we take a new initial approximation $u_0'=u_{N_C}$ and $\Delta_0'=\Delta_{N_C}$ and continue the calculation via algorithm (4) with the halt condition of (7) to get some new number $N_C'\geqslant N_C$, for which $\Delta_0'\geqslant\Delta_{N_C'}>\Delta_e$. We continue this process until condition (7) is met after the first iteration on the latest re-initiation, which gives the best approximation. These arguments show that renewal of the iteration in such a problem should be based not on arguments of matching the number of iterations to the dimensions of the problem, in contrast to correctly formulated problems, but on conditions for matching the error of the input data. Inequalities (8)-(10) can also have practical application in estimating the error of the algorithms, e.g., if we have an estimate for the norm of the solution in space U.

Theorem 3 gives an interpretation of the gradient algorithms of [7, 8]. It can be said that the approximations derived by means of the algorithm of [7] will approximate to the solution on average, because the algorithm was constructed on the assumption that the solution is a function integrable in square. In that respect, successful use of the algorithm requires a knowledge of the magnitude of the desired solution at the right-hand end of the observation range. The further the estimate is from the true value, the worse the approximations will be, although on average they will describe the function well. Results are presented in Fig. 1a from a calculation on a model example with exact specification of the solution at the right-hand end of the interval of observation. Experience with this algorithm confirms that successful solution of an inverse problem requires a considerable reduction in the region containing the solution by means of additional a priori information. For a wide range of boundary-value problems in thermal conduction, this information can concern the smoothness of the solution. The algorithm of [8] was based on the assumption that the first derivative of the solution is integrable in square, and by virtue of Theorem 3 we have that the approximation will approach the solution uniformly. Figure 1b gives an example of a calculation via this algorithm. The algorithm of [8] used additional information on the value of the function at the initial instant, which was derived from conditions for conformity for the boundary and initial conditions for the thermal-conduction equation. One can construct algorithms that do not use any other information apart from the smoothness, but this topic falls outside the scope of this paper.

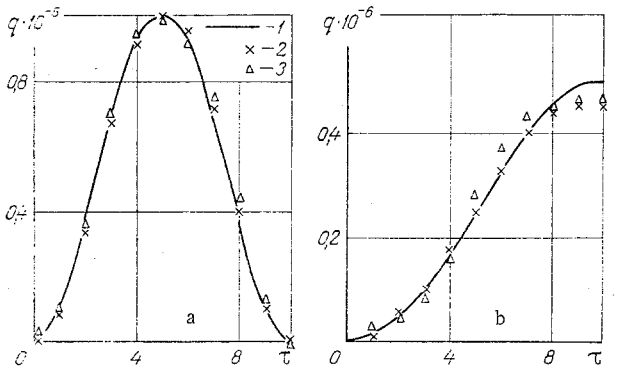

Fig. 1. Recovery of a heat-flux density q(τ), 10^6 W/m² (a by means of the algorithm of [7], b – algorithm of [8]) at one of the boundaries from temperature measurements at the other, which is thermally insulated: 1) true value of the heat flux; 2) value recovered with the temperature specified exactly; 3) same with perturbed data (perturbation up to 10% of the maximum value of the temperature, with a uniform distribution of the perturbation).

Calculations on model examples show that these regularizing gradient algorithms are no less accurate than the classical regularization method (minimization of a smoothing functional) if the a priori information is the same. The algorithms are fast, since 25 iterations can be obtained in the algorithm of [8] with a realization consisting of 50 points in about 4 min of run time for an ES-1020 computer. The required level of discrepancy was attained in the examples in 7-15 iterations with perturbations up to 10% of the maximum value of the right side. The errors in the approximation were considered as small by comparison with the errors in the right sides and were neglected, i.e., the assumption was $\Delta_e = \delta$, for which there is a certain justification in that there was no loss of stability in calculations with $\delta = 0$ (i.e., when there were only errors of approximation), while there was instability even with very small random perturbations in the right sides.

Gradient algorithms allow of simple implementation, and they can be used to advantage in hybrid systems as well as digital ones. All of these advantages of regularizing gradient algorithms indicate that they are very promising for use in automatic systems.

NOTATION

A, A_h, linear operators; U, F, Hilbert spaces; A^{-1}, inverse operator; h, approximation uncertainty; δ, rhs uncertainty, U_j, multiplicity of solutions to the reference equation; J'_u, discrepancy gradient; Δ_n, discrepancy for the n-th iteration; β, step of descent for the n-th iteration; p_n, conjugate direction for the n-th iteration; A^*, conjugate operator; $R(A)$, image of the operator A; $N(A)$, kernel of the operator A; Δ_e, discrepancy level determined by the reference data uncertainty.

LITERATURE CITED

1. A. N. Tikhonov and V. Ya. Arsenin, Methods of Solving Incorrectly Formulated Problems [in Russian], Nauka, Moscow (1978).
2. V. M. Fridman, "The convergence of methods of steepest-descent type," Usp. Mat. Nauk, 17, No. 3, 201-208 (1962).
3. M. Z. Nashed, "Steepest descent for singular linear operator equations," SIAM J. Numer. Anal., 7, No. 2, 358-362 (1970).
4. W. J. Kammerer and M. Z. Nashed, "On the convergence of conjugate gradient methods for singular linear operator equations," SIAM J. Numer. Anal., 9, No. 2, 165-181 (1972).
5. S. F. Gilyazov, "Iterative methods of solving incompatible equations with non-self-conjugate operators," Vestn. Mosk. Univ., Series 15, No. 4, 12-19 (1977).
6. V. A. Morozov, "Regularizing families of operators," in: Computational Methods and Programming [in Russian], Issue 8, Moscow State Univ., Moscow (1967), pp. 63-95.
7. O. M. Alifanov, "Solution of an inverse problem in thermal conduction by iterative methods," Inzh.-Fiz. Zh., 26, No. 4, 682-689 (1974).
8. O. M. Alifanov and S. V. Rumyantsev, "A method of solving incorrectly formulated problems," Inzh.-Fiz. Zh., 34, No. 2, 328-331 (1978).