A GRADIENT METHOD OF DEFINING SMOOTH SOLUTIONS TO INVERSE BOUNDARY-VALUE PROBLEMS IN THERMAL CONDUCTION

E. A. Artyukhin and S. V. Rumyantsev UDC 536.24.02

An iterative algorithm is described for solving boundary-value inverse problems in thermal conduction by steepest descent, which utilizes information on the smoothness of the solution.

Gradient minimization methods are highly efficient in solving inverse problems in thermal conduction [1, 2]. This concerns algorithms that provide mean-square convergence to the solution. An algorithm has been suggested [3] that provides uniform convergence, but in order to use this one needs to know the value of the desired function at one of the points in the observation interval.

Here we provide a generalization of these algorithms as a way of producing solutions that does not require quantitative a priori information about the solution.

The boundary-value problem involves solving an equation of the first kind:

$$Au = f_\delta, \tag{1}$$

where $A: U = L_2[0, \tau_m] \to L_2[0, \tau_m] = F$ is a linear continuous operator that specifies the dependence of the thermal conditions within the body on the unknown boundary condition; $f_\delta = \bar{f} + \tilde{f}$, where $\bar{f}$ is the exact right side and $\tilde{f}$ is the noise, while $\|\tilde{f}\|_{L_2} = \delta$; the inverse operator A^{-1} as a rule is unbounded.

We assume that it is known that the solution u to (1) on the exact data $\bar{f}$ belongs to the space $U_0 = W_2^k[0, \tau_m] \subset U$ (the k-th general derivative exists). In that case we can transfer to the equivalent problem of determining the k-th derivative of the solution and k constants, which are values of the solution and of the derivatives at certain points in the interval of observation, where the resulting approximations will have the required smoothness. To perform the conversion we introduce an auxiliary space $G = L_2[0, \tau_m] \times R^k$ with the elements $g = \{v, c_1, \ldots, c_k\}$ and norm $\|g\|_G^2 = \|v\|_{L_2}^2 + \sum_{i=1}^{k} \rho_i c_i^2$, where $v(\tau) = u^{(k)}(\tau)$, $c_i = u^{(i-1)}(\tau)$, and ρ_i are weights that incorporate the effects of each of the constants c_i.

The function $u \in U_0$ is defined from g in a unique fashion by means of the linear compact operator L: $G \to U_0$:

$$u = Lg = \int_{t_1'}^{t} \ldots \int_{t_k'}^{\tau} v(\xi)\, d\xi d\tau \ldots + \sum_{i=1}^{k} c_i \varphi_i = L_0 v + \sum_{i=1}^{k} c_i \varphi_i,$$

where $\varphi_i(t) = \int_{t_1'}^{t} \ldots \int_{t_{i-1}'}^{\tau} d\xi d\tau$ are polynomials of degree i − 1 and $\varphi_1(t) \equiv 1$; in order to determine u unambiguously, the constants c_i may be found in other ways, but this will influence only the form of φ_i.

Instead of (1) we solve the equivalent equation

$$Bg = f_\delta, \quad B = AL. \tag{2}$$

The sequence in the steepest-descent method for (2) takes the form

$$g_{n+1} = g_n - \beta_n J' g_n = g_n - \beta_n B^* (Bg_n - f_\delta), \quad \beta_n = \frac{\|J' g_n\|_G^2}{\|BJ' g_n\|_{L_2}^2}, \tag{3}$$

where $B^* = L^* A^*$ is an operator conjugate to operator B and $J(g) = \frac{1}{2} \|Bg - f_\delta\|_{L_2}^2$ is the discrepancy functional.

We apply L to both parts of (3) to obtain the transform in space U and to define at once the approximations u_n to $\bar{u}$ without reference to the g_n:

Translated from Inzhenerno-Fizicheskii Zhurnal, Vol. 39, No. 2, pp. 259-263, August, 1980. Original article submitted January 2, 1980.

 0022-0841/80/3902-0862 $07.50

$$u_{n+1} = u_n - \beta_n LJ'g_n = u_n - \beta_n LL^*A^*(Au_n - f_\delta), \quad u_0 = Lg_0,$$

$$\beta_n = \frac{\|L^*A^*(Au_n - f_\delta)\|_G^2}{\|ALL^*A^*(Au_n - f_\delta)\|_{L_2}^2}. \tag{4}$$

The iteration of (4) is halted by reference to the discrepancy, i.e., from the condition $\|Au_n - f_\delta\|_{L_2} \simeq \delta$.

The value of A is calculated from the solution to the conjugate problem for the equation of thermal conduction [4]; we derive an expression for L. From the definition of the conjugate operator,

$$(u, Lg)_{L_2} = (L^*u, g)_G = (g^*, g)_G, \quad \forall u \in U, \quad g \in G,$$

we have

$$(u, Lg)_{L_2} = \sum_{i=1}^{k} c_i (u, \varphi_i)_{L_2} + \int_0^{\tau_m} u(t) \int_{t_1}^{t} \dots \int_{t_k}^{\tau} v(\xi)\, d\xi d\tau \dots dt =$$
$$= \sum_{i=1}^{k} c_i (u, \varphi_i)_{L_2} + \left(\int_{t_1}^{t} u(\tau)\, d\tau \int_{t_2}^{t} \dots \int_{t_k}^{\tau} v(\xi)\, d\xi d\tau \dots\right)\Big|_0^{\tau_m} -$$
$$- \int_0^{\tau_m} \left(\int_{t_1}^{t} u(\tau)\, d\tau \int_{t_2}^{t} \dots \int_{t_k}^{\tau} v(\xi)\, d\xi d\tau \dots\right) dt = \sum_{i=1}^{k} c_i (u, \varphi_i)_{L_2} +$$
$$+ \int_0^{\tau_m} \left(L_1^* u \int_{t_2}^{t} \dots \int_{t_k}^{\tau} v(\xi)\, d\xi d\tau \dots\right) dt = \dots = \sum_{i=1}^{k} c_i (u, \varphi_i)_{L_2} +$$
$$+ \int_0^{\tau_m} v(t) L_k^* \dots L_1^* u dt = (L^*u, g)_G = \sum_{i=1}^{k} \rho_i c_i c_i^* + \int_0^{\tau_m} v v^* dt,$$

where

$$L_i^* z = \begin{cases} \int_t^{\tau_m} z(\tau)\, d\tau, & t > t_i, \\ \int_0^{t} z(\tau)\, d\tau, & t < t_i. \end{cases}$$

Therefore, we have

$$L^*u = \left\{L_k^* \dots L_1^* u, \frac{(u, \varphi_1)_{L_2}}{\rho_1}, \dots, \frac{(u, \varphi_k)_{L_2}}{\rho_k}\right\}.$$

Then the sequence of (4) takes the form

$$u_{n+1} = u_n - \beta_n \left[\sum_{i=1}^{k} \frac{(A^*(Au_n - f_\delta), \varphi_i)_{L_2}}{\rho_i} + \int_{t_1}^{t} \dots \int_{t_k}^{\tau} (L_k^* \dots \right.$$
$$\left. \dots L_1^* A^*(Au_n - f_\delta))\, d\xi d\tau \dots\right],$$
$$u_0 = Lg_0.$$

Transfer to the auxiliary problem of (2) is necessary in order to choose the direction of descent in the initial space U, since this must not cause the sequence to deviate from set U_0 that contains the solution u to (1) on the exact data. The initial approximation u_0 should belong to U_0, and in particular we can take $u_0(\tau) \equiv 0$.

Complete use of the information on the smoothness improves the approximations considerably; a previous method [1] of steepest descent for (1) involved choosing the halt number N(δ) from the discrepancy, which gave mean-square convergence. This fact is not proved here, but numerous model examples and certain theoretical studies [5, 6] point to this conclusion. On the other hand, this algorithm is that of (3) for the auxiliary problem. Therefore, we can assume that the solutions $g_{N(\delta)}$ to the auxiliary problem will converge to $\bar{g} = L^{-1}\bar{u}$ for $\delta \to 0$ in the norm of space G. As L is continuous in the mapping from $L_2[0, \tau_m]$ into $C^k[0, \tau_m]$, we have uniform convergence of the transforms $u_{N(\delta)} = Lg_{N(\delta)}$ to $\bar{u}$ with their derivatives up to order k − 1 inclusive. This conclusion is confirmed by calculations on model examples. The result from one calculation is given below.

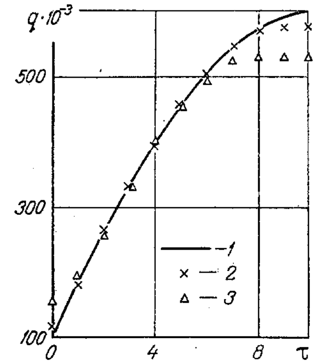

Fig. 1. Recovery of a heat-flux density (q is heat flux in 10^6 W/m^2 and τ is time in sec): 1) exact solution; 2) flux recovered from exact data; 3) the same from corrupted data (the perturbations are uniformly distributed with a maximum spread of 10% of the maximum value of the exact data).

This algorithm is particularly effective if one has a priori information on the values of the solution or of the derivatives at some points in the interval of observation, e.g., values of the temperature of the heated surface at the start, the positions of turning points, or those of points of inflection. In that case, some of the constants c_j are known, and the corresponding functions φ_j may be put as zero, which reduces the volume of computation and improves the approximation. The initial approximation should be chosen from the same class as the desired solution, e.g., $u_0 = \sum_j c_j \varphi_j$, where the summation is carried over the known constants c_j.

As a model example we consider recovery of the heat flux to a thermally insulated rod with constant thermophysical characteristics by reference to measurement of the temperature at the internal end. This involves solving an integral equation of the first kind:

$$\int_0^\tau K(\tau-\xi)\, q(\xi)\, d\xi = T_\delta(\tau), \tag{5}$$

where $q(\xi)$ is the heat flux and $T_\delta(\tau)$ is the measured temperature. It is assumed that $q(\xi)$ has a piecewise-continuous first derivative. The value $q(0)$ is taken as the unknown constant c. The sequence of (4) for this case takes the form

$$q_{n+1}(\tau) = q_n(\tau) - \beta_n \left[\frac{1}{\rho} \int_0^{\tau_m} (T_n(\tau) - T_\delta(\tau))\, d\tau + \right.$$

$$\left. + \int_0^\tau \int_t^{\tau_m} \int_\xi^{\tau_m} K(\eta-\xi)\, (T_n(\eta) - (T_\delta(\eta))\, d\eta d\xi dt \right] = q_n(\tau) - \beta_n \Delta q_n(\tau),$$

$$\beta_n = \left\{\frac{1}{\rho}\left[\int_0^{\tau_m} (T_n(\tau) - T_\delta(\tau)) d\tau\right]^2 + \int_0^{\tau_m}\left[\int_\tau^{\tau_m}\int_\xi^{\tau_m} K(\eta-\xi)\,(T_n(\eta) - T_\delta(\eta)) d\eta d\xi\right]^2 d\tau\right\}\left\{\int_0^{\tau_m}\left[\int_0^\tau K(\tau-\xi)\,\Delta q_n(\xi)\, d\xi\right]^2 d\tau\right\}^{-1},$$

where

$$T_n(\tau) = \int_0^\tau K(\tau-\xi)\, q_n(\xi)\, d\xi.$$

Figure 1 shows the results.

The weights ρ_i should be chosen such that the effects on the function from each of the unknown constants $\bar{c}_i$ and derivative $\bar{v}$ should be approximately the same, i.e., the norms of the transformations for each of the c_i and v should be consistent. The norm of the c_i transformation is $K_i = \|\varphi_i\|_F / \sqrt{\rho_i}$, while that for the v transformation is $K_v = \|L_0\|$ and then ρ_i may be chosen, for example, from the condition $K_i = (1/i)K_0$, and in that case $\rho_i = i^2 \|\varphi_i\|_F / \|L_0\|^2$. In our model example, $\|\varphi\|_F^2 = \tau_m$, $\|L_0\| = \tau_m/\sqrt{2}$, and therefore $\rho = 2/\tau_m$.

Calculations on model examples show that the choice of the ρ_i substantially influences the convergence rate in the algorithm, but the rate is acceptable if the ρ_i are chosen from the condition for consistency of the transformation norms.

Similar arguments allow one to construct an algorithm for the conjugate-gradient method, which is of somewhat higher performance with virtually the same run time.

NOTATION

A, B, linear operators; u, element of solution space U; $\bar{f}$, exact reference data; $\tilde{f}$, reference data uncertainty; δ, value of reference data uncertainty; A^{-1}, inverse operator; $u^{(k)}(\tau)$, k-th derivative of function u; τ_m, length of observation interval; $\varphi_i(t)$, polynomials of degree $i-1$; A^*, B^*, L^*, operators conjugate to the operators A, B, L; $J'g$, discrepancy functional gradient; β_n, descent step along the discrepancy antigradient for the n-th iteration; $K(\tau-\xi)$, kernel of integral equation; $q(\tau)$, heat flux; $T_\delta(\tau)$, measured temperature inside body.

LITERATURE CITED

1. O. M. Alifanov, "Solution of an inverse problem in thermal conduction by iterative methods," Inzh.-Fiz. Zh., 26, No. 4, 682-689 (1974).
2. O. M. Alifanov, "Determination of heat loads from a solution to a nonlinear inverse problem," Teplofiz. Vys. Temp., 15, No. 3, 598-605 (1977).
3. O. M. Alifanov and S. V. Rumyantsev, "A method of solving incorrectly formulated problems," Inzh.-Fiz. Zh., 34, No. 2, 328-331 (1978).
4. B. M. Budak and F. P. Vasil'ev, Approximate Methods of Solving Optimum-Control Problems [in Russian], Issue 2, Moscow State Univ., Moscow (1969).
5. V. M. Fridman, "Convergence of methods of steepest-descent type," Usp. Mat. Nauk, 17, No. 3, 201-208 (1973).
6. M. Z. Nashed, "Steepest descent for singular linear operator equations," SIAM J. Numer. Anal., 7, No. 2, 358-362 (1970).

OPTIMAL CHOICE OF DESCENT STEPS IN GRADIENT METHODS OF SOLUTION OF INVERSE HEAT-CONDUCTION PROBLEMS

E. A. Artyukhin and S. V. Rumyantsev

UDC 536.24.02

Modifications are proposed for the methods of steepest descent and conjugate gradients for the solution of multiparameter inverse problems in heat conduction.

In the solution of inverse heat-conduction problems it often becomes necessary to determine several independent functions and parameters at once. Such multiparameter problems arise in the solution of coefficient-type inverse problems, in the joint determination of the external thermal load and some thermophysical characteristics of the body, etc. An attempt to take the most complete account of the a priori information about the desired solution may also lead to such problems.

In the solution of boundary-value inverse problems with one unknown (a function or a parameter) it has been found very effective to use algorithms based on gradient methods of minimization [1-3]. The use of these methods in a case when it is necessary to determine several independent variables is made more difficult by the fact that the descent step is chosen to be the same for all components of the direction of descent. Such a method of choosing the step frequently leads to very slow convergence of the gradient methods. The convergence may be speeded up considerably by choosing different descent steps for the different components of the gradient of the minimizing functional, i.e., to determine not one step but a vector of steps from the condition that the target functional has a minimum with respect to this vector at each iteration.

We shall show how this method can be used for constructing gradient algorithms for the solution of boundary-value inverse problems in heat conduction when a priori information concerning the smoothness of the desired solution is available.

A boundary-value inverse heat-conduction problem for bodies with constant thermophysical characteristics can be reduced to the solution of the first-order equation

$$Au = f_\delta, \quad u \in U, \quad f_\delta \in F, \tag{1}$$

Translated from Inzhenerno-Fizicheskii Zhurnal, Vol. 39, No. 2, pp. 264-269, August, 1980. Original article submitted January 2, 1980.