M. H. F A R A G (Minia)

THE GRADIENT PROJECTION METHOD FOR SOLVING AN OPTIMAL CONTROL PROBLEM

Abstract. A gradient method for solving an optimal control problem described by a parabolic equation is considered. The gradient projection method is applied to solve the problem. The convergence of the projection algorithm is investigated.

1. Introduction. The theory of optimal control systems with distributed parameters is one of the leading sections of optimization theory. It has wide applications in various practical fields. The theory of optimal control problems has been studied by many workers [1, 2, 6, 7]. They have shown [4, 9, 10] that these problems arise in many physical applications such as heat conductivity, filtration and diffusion.

2. Statement of the problem and definitions. Let it be required to minimize the function

$$(1) \qquad f(v) = \int_0^l |u(x,T;v) - g(x)|^2 \, dx + \beta \int_0^T |v_1(t)|^2 \, dt$$

provided that $u(x,t;v)$ is a solution of the boundary value problem

$$(2) \quad u_t = a^2 u_{xx} + B(x,t)u + v_2(x,t), \quad (x,t) \in \Omega = [0 < x < l, 0 < t \le T],$$

$$(3) \qquad\qquad u(x,0) = \phi(x), \qquad 0 \le x \le l,$$

$$(4) \qquad u_x(0,t) = 0, \quad u_x(l,t) = \nu[v_1(t) - u(l,t)], \quad 0 < t \le T,$$

where a^2, l, ν, T, β are positive numbers, $v_1(t)$ the temperature of the external medium, $v_2(x,t)$ the density of heat sources, and the control v is in

$$V = \left\{ v \mid v = (v_1(t), v_2(x,t)); v_1(t) \in L_2[0,T], v_{1\min} \leq v_1(t) \leq v_{1\max}; \right.$$
$$\left. v_2(x,t) \in L_2(\Omega), \int_0^l \int_0^T |v_2(x,t)|^2 \, dx \, dt \leq R^2 \right\},$$

where $v_{1\min} < v_{1\max}; R > 0$ is a given number; $g(x), \phi(x) \in L_2[0,l], B(x,t) \in L_2(\Omega)$ are given functions and $H = L_2[0,T] \times L_2(\Omega)$.

DEFINITION 1. The problem of finding a function $u = u(x,t;v)$ satisfying conditions (2)–(4) for a given $v \in V$ is called the *reduced problem*.

DEFINITION 2. The *solution* of the reduced problem (2)–(4) *corresponding to* $v \in V$ is a function $u(x,t) \in H^{1,0}(\Omega(4))$ satisfying the integral identity

$$(5) \quad \int_0^l \int_0^T [-u\eta_t + a^2 u_x \eta_x + B(x,t)u\eta - v_2(x,t)\eta] \, dx \, dt$$
$$= \int_0^l \phi(x)\eta(x,0) \, dx + a^2 \nu \int_0^T [v_1(t) - u(l,t)]\eta(l,t) \, dt$$

for all $\eta = \eta(x,t) \in H^1(\Omega)$ with $\eta(x,T) = 0$.

Equations (1)–(4) are the mathematical formulation of the optimal control problem for a linear parabolic equation with controls in boundary conditions and the right side of equation (2). Optimal control problems for linear and nonlinear parbolic equations have been widely considered in the literature (see for instance [4, 8, 18]), and were studied by Madatov [11] and Mokrane [12], where the existence, uniqueness and regularity of the solution were proved. In addition, Farag [3] and Phillipson and Mitter [13] have derived numerical results for the heat equation with strong nonlinearity.

3. The gradient of the function. The principal result in this section is Theorem 3.1. Its proof will be prepared by two lemmas:

LEMMA 3.1. *Let $\delta u(x,t)$ be the generalized solution of the boundary value problem*

$$(6) \qquad \delta u_t - a^2 \delta u_{xx} - B(x,t)\delta u - \delta v_2 \ (x,t) = 0, \ (x,t) \in \Omega,$$
$$(7) \qquad \qquad \delta u(x,0) = 0, \qquad 0 \leq x \leq l,$$
$$(8) \qquad \delta u_x(0,t) = 0, \quad \delta u_x(l,t) = \nu[\delta v_1(t) - \delta u(l,t)], \quad 0 < t \leq T.$$

Then

$$(9) \qquad \int_0^l |\delta u(x,T)|^2 \, dx \leq C \left[\int_0^T |\delta v_1(t)|^2 \, dt + \int_0^l \int_0^T |\delta v_2(x,t)|^2 \, dx \, dt \right]$$
$$= C \|\delta v\|_H^2,$$

where $C > 0$ is a constant which is independent of the choice of $\delta v \in V$.

Proof. We multiply (6) by δu and integrate it on the rectangle Ω. By using the conditions (7) and (8), we obtain the reduced equation:

$$(10) \quad \frac{1}{2}\int_0^l |\delta u(x,T)|^2\,dx + a^2\nu\int_0^T |\delta u(l,t)|^2\,dt + a^2\int_0^l\int_0^T |\delta u_x|^2\,dx\,dt$$

$$= a^2\nu\int_0^T \delta u(l,t)\delta v_1(t)\,dt + \int_0^l\int_0^T \delta u\delta v_2\,dx\,dt.$$

Applying the inequality $ab \le \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2$, $\varepsilon > 0$, we obtain

$$(11) \quad \frac{1}{2}\int_0^l |\delta u(x,T)|^2\,dx + a^2\nu\int_0^T |\delta u(l,t)|^2\,dt + a^2\int_0^l\int_0^T |\delta u_x|^2\,dx\,dt$$

$$\le \frac{1}{2}a^2\varepsilon_1\nu\int_0^T |\delta u(l,t)|^2\,dt + \frac{1}{2\varepsilon_1}a^2\nu\int_0^T |\delta v_l(t)|^2\,dt$$

$$+ \frac{\varepsilon_2}{2}\int_0^l\int_0^T |\delta u(x,t)|^2\,dx\,dt + \frac{1}{2\varepsilon_2}\int_0^l\int_0^T |\delta v_2(x,t)|^2\,dx\,dt.$$

Since

$$|\delta u(x,t)|^2 = \left(\int_x^l \delta u_x(\theta,t)\,d\theta - \delta u(l,t)\right)^2 \le 2\left(\int_x^l \delta u_x(\theta,t)\,d\theta\right)^2 + 2|\delta u(l,t)|^2$$

$$\le 2l\int_0^l |\delta u_x(x,t)|^2\,dx + 2|\delta u(l,t)|^2$$

we have

$$(12) \quad \int_0^l\int_0^T |\delta u(x,t)|^2\,dx\,dt \le 2l^2\int_0^l\int_0^T |\delta u_x|^2\,dx\,dt + 2l\int_0^T |\delta u(l,t)|^2\,dt.$$

From (11), (12) and by reducing these terms we obtain

$$(13) \quad \frac{1}{2}\int_0^l |\delta u(x,T)|^2\,dx + \left(a^2\nu - \frac{a^2\nu\varepsilon_1}{2} - l\varepsilon_2\right)\int_0^T |\delta u(l,t)|^2\,dt$$

$$+ (a^2 - l^2\varepsilon_2)\int_0^l\int_0^T |\delta u_x|^2\,dx\,dt$$

$$\le \frac{a^2\nu}{2\varepsilon_1}\int_0^T |\delta v_l(t)|^2\,dt + \frac{1}{2\varepsilon_2}\int_0^l\int_0^T |\delta v_2(x,t)|^2\,dx\,dt.$$

Letting $\varepsilon_2 = a^2\varepsilon_1$ and $0 < \varepsilon_1 < \min[1/l^2; 2\nu/(\nu+2l)]$, from (13) we obtain (9) with $C = \max[a^2\nu/\varepsilon_1; 1/(a^2\varepsilon_1)]$. The lemma is proved.

LEMMA 3.2. *Let* $\lambda(x,t;v) = \lambda(x,t)$ *be the generalized solution of the conjugate boundary value problem*

(14) $$\lambda_t = -a^2\lambda_{xx} - B(x,t)\lambda, \qquad (x,t) \in \Omega,$$

(15) $$\lambda(x,T) = 2[u(x,T;v) - g(x)], \qquad 0 \le x \le l,$$

(16) $$\lambda_x(0,t) = 0, \quad \lambda_x(l,t) = -\nu\lambda(l,t), \quad 0 < t < T.$$

Then

(17) $$2\int_0^l [u(x,T;v) - g(x)]\delta u(x,T)\,dx$$
$$= \int_0^T a^2\nu\lambda(l,t;v)\delta v_1(t)\,dt + \int_0^l\int_0^T \lambda(x,t;v)\delta v_2(x,t)\,dx\,dt.$$

P r o o f. Applying the conditions (6)–(8) and (14)–(16), we obtain

(18) $$2\int_0^l [u(x,T,v) - g(x)]\delta u(x,T)\,dx$$
$$= \int_0^l \lambda(x,T)\delta u(x,T)\,dx$$
$$= \int_0^l\int_0^T [\lambda_t\delta u + \lambda\delta u_t]\,dx\,dt$$
$$= \int_0^l\int_0^T [-a^2\lambda_{xx}\delta u + a^2\lambda\delta u_{xx} + \lambda\delta v_2]\,dx\,dt$$
$$= \int_0^T a^2\nu\lambda(l,t;v)\delta v_1(t)\,dt + \int_0^l\int_0^T \lambda(x,t;v)\delta v_2(x,t)\,dx\,dt.$$

The equality (17) is thus obtained. The lemma is proved.

DEFINITION 3. The *solution* of the conjugate boundary value problem (14)–(16) *corresponding to* $v \in V$ is a function $\lambda(x,t) \in H^{1,0}(\Omega)$ satisfying the integral identity

(19) $$\int_0^l\int_0^T [-\lambda\xi_t + a^2\lambda_{xx}\xi + B(x,t)\lambda\xi]\,dx\,dt$$
$$= -2\int_0^l [u(x,T;v) - g(x)]\xi(x,T)\,dx$$

for all $\xi = \xi(x,t) \in H^1(\Omega)$ with $\xi(x,0) = 0$.

THEOREM 3.1. *The function* (1) *is differentiable in* H *and its gradient at* $v \in V$ *is given by*

$$(20) \qquad f_v(v) = \frac{\partial f}{\partial v} = -\frac{\partial \Re}{\partial v} \equiv \left(-\frac{\partial \Re}{\partial v_1}, -\frac{\partial \Re}{\partial v_2} \right),$$

where $\Re$ *is defined by*

$$\Re(x, t, \lambda, v_1, v_2) \equiv -[a^2 \nu v_1 \lambda(l, t; v_1) + \beta v_1^2 + v_2 \lambda(x, t; v_2)].$$

P r o o f. Consider the increment of the function (1):

$$(21) \quad \delta f(v) = f(v + \delta v) - f(v)$$

$$= 2 \int_0^l [u(x, T, v) - g(x)] \delta u(x, T) \, dx + 2\beta \int_0^T v_1(t) \delta v_1(t) \, dt$$

$$+ \int_0^l |\delta u(x, T)|^2 \, dx + \beta \int_0^T |\delta v_1(t)|^2 \, dt$$

where $v \in V$, $v + \delta v \in V$, $\delta u(x, t) \equiv u(x, t; v + \delta v) - u(x, t; v)$, $u \equiv u(x, t; v)$.

By substituting equality (17) and estimate (9) in (21), it follows that the function (1) is differentiable in H and its gradient is given by the expression (20). The theorem is proved.

4. The gradient projection method. One of the first authors who used projection methods for solving constrained problems was J. B. Rosen [16, 17]. A lot of projection algorithms were described by Polak [14] and Pshenichnyĭ and Danilin [15]. Having the gradient function (1), we can use the gradient projection method for solving the problem (1)–(4). According to this method we construct a sequence $\{v^k = (v_1^k(t), v_2^k(x, t))\}$ by setting

$$(22) \qquad v_1^{k+1} = \begin{cases} v_1^k - \gamma^k f_v(v_1^k) & \text{if } v_{1\min} \le Z_1(v_1^k) \le v_{1\max}, \\ v_{1\min} & \text{if } Z_1(v_1^k) < v_{1\min}, \\ v_{1\max} & \text{if } Z_1(v_1^k) > v_{1\max}, \end{cases}$$

$$(23) \qquad v_2^{k+1} = \begin{cases} v_2^k - \gamma^k f_v(v_2^k) & \text{if } Z_2(v_2^k) \le R^2, \\ \dfrac{R[v_2^k - \gamma^k f_v(v_2^k)]}{\sqrt{Z_2(v_2^k)}} & \text{if } Z_2(v_2^k) > R^2, \end{cases}$$

where $Z_1(v_1^k) = v_1^k - \gamma^k f_v(v_1^k)$ and $Z_2(v_2^k) = \int_0^l \int_0^T |v_2^k - \gamma^k f_v(v_2^k)|^2 \, dx \, dt$.

The values $\gamma^k \ge 0$ in (21)–(22) may be selected in one of the following ways:

(i) γ^k is defined by

$$(24) \qquad \gamma^k = \min_{\gamma \ge 0} f(\gamma) = \min_{\gamma \ge 0}(f(v^k) - \gamma f_v(v^k)).$$

(ii) If the gradient $f_v(v)$ satisfies the condition

$$(25) \qquad \|f_v(v) - f_v(w)\|_H \leq L\|v - w\|_H$$

for any $v, w \in V$, $L = \text{const} > 0$, then γ^k may be found from the conditions

$$(26) \qquad 0 < c_1 \leq \gamma^k \leq \frac{2}{L + 2c_2}.$$

Here, c_1, $c_2 > 0$ are parameters selected by computer.

(iii) The parameter $\gamma^k \in [0, 1]$ can be chosen from the monotonicity condition $f(v^{k+1}) < f(v^k)$.

(iv) γ^k can be chosen from the condition

$$(27) \qquad f(v^k) - f(v^k - \gamma^k f_v(v^k)) \geq \varepsilon\gamma^k\|f_v(v^k)\|^2, \qquad \varepsilon > 0.$$

THEOREM 4.1. *Let V be a closed convex subset of H, and $f \in C^{1,1}(V)$ with $f_* = \inf_V f(v) > -\infty$. Let $\{v^k\}$ be the sequence of controls generated by the projection algorithm formulated in (22)–(27) for an arbitrary initial approximation $\{v^0\} \in V$. Then the sequence $\{f(v^k)\}$ decreases and $\lim_{k \to \infty} \|v^k - v^{k+1}\| = 0$. Moreover, if f is convex in H and the set $M(v^0) = \{v^0 \in V : f(v) \leq f(v^0)\}$ is bounded, then the sequence $\{v^k\}$ minimizes the function $f(v)$ in V and converges to v_* weakly in H, and it also satisfies the estimate*

$$(28) \qquad 0 \leq f(v^k) - f_* \leq \frac{c_3}{k}, \qquad k = 1, 2, \ldots; \; c_3 = \text{const} \geq 0.$$

If f is also strongly convex in V, then $\{v^k\}$ converges to the unique minimum control v_* such that

$$(29) \qquad \|v^k - v_*\|^2 \leq \frac{c_4}{k}, \qquad k = 1, 2, \ldots; c_4 = \text{const} \geq 0.$$

The proof directly follows from that of Theorem 5.2.1 of [19].

References

[1] A. G. Butkovskiĭ, *Optimal Control Theory for Systems with Distributed Parameters*, Nauka, Moscow, 1965 (in Russian).

[2] Yn. V. Egorov, *On some optimal control problems*, Zh. Vychisl. Mat. i Mat. Fiz. 3 (1963), 887–904 (in Russian).

[3] M. H. Farag, *A numerical solution to a nonlinear problem of the identification of the characteristics of a mathematical model of heat exchange*, in: *Mathematical Modeling and Automated Systems*, A. D. Iskenderov (ed.), Bakin. Gos. Univ., Baku, 1990, 23–30 (in Russian).

[4] M. H. Farag and S. H. Farag, *An existence and uniqueness theorem for one optimal control problem*, Period. Math. Hungar. 30 (1995), 61–65.

[5] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, Englewood Cliffs, N.J., 1964.

[6] A. D. Iskenderov, *On a certain inverse problem for quasilinear parabolic equations*, Differentsial'nye Uravneniya 10 (1974), 890–898 (in Russian).

[7] A. D. Iskenderov and R. K. Tagiev, *Optimization problems with controls in coefficients of parabolic equations*, ibid. 19 (1983), 1324–1334 (in Russian).

[8] J.-L. Lions, *Control problems in systems described by partial differential equations*, in: *Mathematical Theory of Control*, A. V. Balakrishnan and L. W. Neustadt (eds.), Academic Press, New York and London, 1969, 251–271.

[9] —, *Optimal Control by Systems Described by Partial Differential Equations*, Mir, Moscow, 1972 (in Russian).

[10] K. A. Lurie, *Optimal Control in Problems of Mathematical Physics*, Nauka, Moscow, 1975 (in Russian).

[11] M. D. Madatov, *Regularization of one class of optimal control problems*, in: *Approximate Methods and Computer*, A. D. Iskenderov (ed.), Bakin. Gos. Univ., Baku, 1982, 78–80 (in Russian).

[12] A. Mokrane, *An existence result via penalty method for some nonlinear parabolic unilateral problems*, Boll. Un. Mat. Ital. B 8 (1994), 405–417.

[13] G. A. Phillipson and S. K. Mitter, *Numerical solution of a distributed identification problem via a direct method*, in: *Computing Methods in Optimization Problems—2*, L. A. Zadeh, L. W. Neustadt and A. V. Balakrishnan (eds.), Academic Press, New York, 1969, 305–315.

[14] E. Polak, *Computational Methods in Optimization*, Academic Press, New York, 1971.

[15] B. N. Pshenichnyĭ and Yu. M. Danilin, *Numerical Methods in Extremal Problems*, Mir, Moscow, 1982.

[16] J. B. Rosen, *The gradient projection method for nonlinear programming. Part I*: *Linear constraints*, SIAM J. Appl. Math. 8 (1960), 181–217.

[17] —, *The gradient projection method for nonlinear programming. Part II*: *Nonlinear constraints*, ibid. 9 (1961), 514–532.

[18] Ts. Tsachev, *Optimal control of linear parabolic equation*: *The constrained right-hand side as control function*, Numer. Funct. Anal. Optim. 13 (1992), 369–380.

[19] F. P. Vasil'ev, *Numerical Methods for Solving Extremal Problems*, Nauka, Moscow, 1988 (in Russian).

M. H. Farag
Department of Mathematics
Faculty of Science
Minia University
Minia, Egypt