Kangweon-Kyungki Math. Jour. **6** (1998), No. 1, pp. 27–46

ON OPTIMAL CONTROL OF A BOUNDARY VALUE PROBLEM

Hongchul Kim and Gye-Soo Rim

Abstract. We are concerned with an optimal control problem governed by a Poisson equation in which body force acts like a control parameter. The cost functional to be optimized is taken to represent the error from the desired observation and the cost due to the control. We recast the problem into the mixed formulation to take advantage of the minimax principle for the duality method. The existence of a saddle point for the Lagrangian shall be shown and the optimality system will be derived therein. Finally, to attain an optimal control, we combine the optimality system with an operational technique. By achieving the gradient of the cost functional, a convergent algorithm based on the projected gradient method is established.

1. Introduction

Optimal control problems described by the distributed parameter systems have a variety of mechanical and technical sources and applications. Fundamental class of optimal controls and its mathematical approaches can be found in Lions[6]. In this paper, we will examine a thorough mechanism for an optimal control problem governed by the elliptic system through two canonical approaches: one by operational approach and one by duality method. Our purpose is to give a precise mathematical mechanism to solve the optimal control problem governed by the distributed parameter system.

Let us describe our model problem. Let Ω be a bounded open set in $\mathbb{R}^n$ whose boundary $\partial\Omega$ is a $\mathcal{C}^1$-manifold with Ω on one side. For a

Received December 23, 1997.

1991 Mathematics Subject Classification: 35J25, 49J20, 65K10.

Key words and phrases: optimal control, duality method, minimax principle, optimality system, projected gradient method.

given $f \in \mathrm{H}^{-1}(\Omega)$, we are concerned with the Poisson equation

$$\begin{cases} -\Delta u = f + g \quad \text{in } \Omega\,, \\ \quad u = 0 \quad \text{on } \partial\Omega\,, \end{cases} \tag{1.1}$$

where g is a control variable through the body. One may find applications of this problem in elastic membranes, electrostatics, fluid flow, steady-state heat conduction and many other topics and physical situations. One may regard g as a heat flux over the body for example. Our objective is to find an optimal heat flux g through the body closest to a given state $\omega \in \mathrm{L}^2(\Omega)$. Since each control g is exerted at some cost, we are concerned with the minimization of the functional

$$J(u,g) = \frac{1}{2}\int_\Omega (u-\omega)^2\,dx + \frac{\alpha}{2}\int_\Omega (g)^2\,dx\,, \tag{1.2}$$

over some given set of admissible controls $g \in \mathfrak{U}_{ad} \subset \mathrm{L}^2(\Omega)$. Here, the first term of (1.2) represents error of the heat from the desired observation and the second one the cost due to the control. Our aim is to find an optimal heat flux g through the body to minimize the functional J. We shall assume that α is a positive constant. This assumption is essential in our environment of the problem. With regard to this, see the remark after section 3.1.

1.1. Notations and Preliminaries

We denote by $\mathrm{H}^s(\Omega)$, the standard Sobolev space of order s with respect to the domain Ω. The norm on $\mathrm{H}^s(\Omega)$ is denoted by $\|\cdot\|_s$. Especially, when $s = 0$, $\|\cdot\|_0$ represents $\mathrm{L}^2(\Omega)$-norm. For a sake of brevity, we will denote $\|\cdot\|_0$ by $\|\cdot\|$. For the space of interest to us, we define the space over which the homogeneous boundary conditions are imposed by

$$\mathrm{H}^1_0(\Omega) = \{\, v \in \mathrm{H}^1(\Omega) \mid v = 0 \ \text{ on } \ \partial\Omega\}\,.$$

The seminorm $|\cdot|$ on $\mathrm{H}^1(\Omega)$ defined by

$$|v| = \Big(\int_\Omega \nabla v \cdot \nabla v\,dx\Big)^{1/2} = \Big(\int_\Omega \sum_{i=1}^n (\frac{\partial v}{\partial x_i})^2\,dx\Big)^{1/2}$$

is equivalent to the norm $\|\cdot\|_1$ on $\mathrm{H}^1(\Omega)$ by Poincare's lemma([3],[4]), i.e., there exists a positive constant c independent of Ω and $v \in \mathrm{H}_0^1(\Omega)$ such that

$$|v| \geq c\|v\|_1\,, \quad \forall v \in \mathrm{H}_0^1(\Omega)\,. \tag{1.3}$$

We define the inner product on $\mathrm{H}^s(\Omega)$ by $(\cdot,\cdot)_s$ and $(\cdot,\cdot)_0$ by $(\cdot,\cdot)$. By $< \cdot,\cdot >_s$, we denote the duality pairing between $\mathrm{H}_0^s(\Omega) \equiv \mathrm{H}^s(\Omega) \cap \mathrm{H}_0^1(\Omega)$ and its dual $\mathrm{H}^{-s}(\Omega)$. Especially, when $s = 1$, we will denote the duality between $\mathrm{H}_0^1(\Omega)$ and its dual $\mathrm{H}^{-1}(\Omega)$ by $< \cdot,\cdot >$ for simplicity. When X and Y are Banach spaces, we denote the class of bounded linear operators from X into Y by $\mathcal{L}(X,Y)$. We also denote the continuous(or bounded) linear functional on a Hilbert space V by V^*.

To deal with the derivation of functional, we are concerned with the Gateaux-differential which is defined by the following: Let X and Y be normed spaces and $T : U \subset X \longrightarrow Y$ be a mapping of an open subset U of X into Y. By $T'(x;\phi)$, we shall represent the Gateaux-derivative(or simply, G-derivative) of T at $x \in U$ in the direction $\phi \in X$. Especially, when X is a Hilbert space and $Y = \mathbb{R}$, if $T'(x;\cdot)$ belongs to X^*, there exists a unique element $\mathcal{G}(x) \in X^*$ such that

$$T'(x;\phi) = < \mathcal{G}(x), \phi >_{X^* \times X} \quad \text{for all } \phi \in X$$

by Riesz representation theorem([8]). This $\mathcal{G}(x) \in X^*$ is called the *gradient* of T at $x \in U$ and is often denoted by $\nabla T(x)$. In connection with the weak variational formulation of (1.1), the following result plays an essential role.

LEMMA 1.1. (**Lax-Milgram Lemma**) Let V be a real separable Hilbert space equipped with the norm $\|\cdot\|_V$. Let $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ be a bilinear form satisfying

$$|a(u,v)| \leq c_1\|u\|_V\|v\|_V\,, \quad \forall u,v \in V\,, \quad (V\text{-}continuity) \tag{1.4}$$

and

$$a(v,v) \geq c_2\|v\|_V^2\,, \quad \forall v \in V\,, \quad (V\text{-}coercivity) \tag{1.5}$$

where c_1 and c_2 are positive constants independent of Ω and u, $v \in V$. Then, for each $f \in V^*$, there exists a unique $u \in V$ satisfying

$$a(u,v) =< f, v >, \quad \forall v \in V$$

and

$$\|u\|_V \leq \frac{1}{c_2} \|f\|_{V^*} .$$

For the proof, one may refer to [3], [4] and [8].

1.2. Weak Formulation of the Problem

Throughout this paper, we will denote $V = \mathrm{H}_0^1(\Omega)$. As the bilinear form over $V \times V$, we will define

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx = (\nabla u, \nabla v) .$$

Let us assume that the admissible set $\mathfrak{U}_{ad}$ is a nonempty closed convex subset of $\mathrm{L}^2(\Omega)$. Since the inclusions of $V \subset \mathrm{L}^2(\Omega) \subset V^*$ are continuous embeddings, for a given $g \in \mathfrak{U}_{ad}$, one can pose the weak variational formulation of the equations (1.1) into

$$a(u,v) =< f, v > +(g,v) , \quad \forall v \in V . \tag{1.6}$$

LEMMA 1.2. *For a given $g \in \mathfrak{U}_{ad}$, the weak problem (1.6) is well-posed, i.e., for a given $g \in \mathfrak{U}_{ad}$, there exists a unique solution u which satisfies (1.6) and depends merely on the data f and g.*

Proof. The bilinear form $a(\cdot,\cdot)$ is V-continuous and V-coercive by (1.3). Hence, by Lax-Milgram Lemma, given $g \in \mathfrak{U}_{ad}$ there exists a unique solution u of (1.6) such that

$$\|u\|_1 \leq c(\|f\|_{-1} + \|g\|) ,$$

for some positive constant c independent of the data. □

Since this solution u depends on the control parameter $g \in \mathfrak{U}_{ad}$, one can write $u = u(g)$ by regarding u as a function $u : \mathfrak{U}_{ad} \rightarrow V$. If we put

$w = u(g_1) - u(g_2)$ for any two $g_1,\ g_2 \in \mathfrak{U}_{ad}$, then since w is a solution of

$$\begin{cases} -\Delta w = g_1 - g_2 \quad \text{in } \Omega \\ \quad w = 0 \quad \text{on } \partial\Omega\,, \end{cases}$$

we have

$$\|w\|_2 = \|u(g_1) - u(g_2)\|_2 \le \kappa_1 \|g_1 - g_2\| \tag{1.7}$$

for some positive constant κ_1 by the canonical regularity of the elliptic system([4], [7]). So, $u : \mathfrak{U}_{ad} \to V$ is injective in a sense.

Based on these facts, one can set the functional (1.2) to be optimized as a functional over $\mathfrak{U}_{ad}$ which is defined by

$$J(g) = J(u(g), g) = \frac{1}{2}\|u(g) - \omega\|^2 + \frac{\alpha}{2}\|g\|^2\,. \tag{1.8}$$

Then, our problem is reduced to the problem of seeking $g^* \in \mathfrak{U}_{ad}$ such that

$$J(g^*) = \inf_{g \in \mathfrak{U}_{ad}} J(g)\,. \tag{1.9}$$

If such a g^* exists, it is called an *optimal control* for the problem.

2. Derivation of the Optimality System

In this section, we derive the optimality system to find an optimal control. The major role of optimality system is to obtain the gradient of the functional, from which one can set up a scheme to approximate the optimal control. We recast our problem into a mixed formulation by employing the duality method of the saddle point formulation of the problem. The existence of an optimal control will be shown as natural results.

2.1. Duality Method

Ahead of our exposition, we need the following conventions and facts. A function $T : U \subset X \longrightarrow X^*$ on a nonempty subset U of a normed space X into its dual is called *monotone* if it satisfies the following inequality

$$< Tx - Ty, x - y >_{X^* \times X} \geq 0 \,, \quad \forall x, y \in U \,.$$

We call T to be a *coercive* operator if

$$\lim_{\|x\|_X \to \infty} \Big(\frac{< Tx, x >}{\|x\|_X} \Big) = +\infty \,.$$

Our approach is utterly based on the following fundamental result.

Lemma 2.1. ([8]) Let K be a nonempty closed convex subset of a separable Hilbert space X and let $T : K \to \mathbb{R}$ be G-differentiable on K. Assume the gradient ∇T is monotone and coercive, then the set $M = \{x \in K : T(x) \leq T(y) \quad \forall y \in K\}$ is a nonempty closed convex set and $x \in M$ if and only if

$$x \in K \text{ and } T'(x; y - x) \geq 0 \,, \quad \forall y \in K \,.$$

Moreover, if ∇T is strictly monotone, M contains at most one point.

We make use of the saddle point approach to derive the existence of the optimal solution and the optimality system needed. For this purpose, we first describe some basic convention and related facts. For details, one may consult [2] or [5].

Let X,Y be two Banach spaces and U,Λ nonempty, convex, closed subsets of X,Y, respectively. Let $\mathcal{F} : U \times \Lambda \to \mathbb{R}$ be a function. We say that $(u, \lambda) \in U \times \Lambda$ is a *saddle point* of $\mathcal{F}$ on $U \times \Lambda$ if and only if

$$\mathcal{F}(u, \mu) \leq \mathcal{F}(u, \lambda) \leq \mathcal{F}(v, \lambda) \,, \quad \forall v \in U \,, \ \forall \mu \in \Lambda \,.$$

Concerning the existence and eventually the uniqueness of the saddle point, we have the following results.

Theorem 2.2. ([5]) *Let*

[i] *For every $v \in U$ the function $\mathcal{F}(v, \cdot)$ is concave and upper-semicontinuous on Λ;*

[ii] *For every $\mu \in \Lambda$ the function $\mathcal{F}(\cdot, \mu)$ is convex and lower-semicontinuous on U;*

[iii] *There exists* $\mu_0 \in \Lambda$ *such that*

$$\lim_{\substack{\|v\|_X \to \infty \\ v \in U}} \mathcal{F}(v, \mu_0) = +\infty\,;$$

[iv] *There exists* $v_0 \in U$ *such that*

$$\lim_{\substack{\|\mu\|_Y \to \infty \\ \mu \in \Lambda}} \mathcal{F}(v_0, \mu) = -\infty\,.$$

Then, there exists at least one saddle point of $\mathcal{F}$ *on* $U \times \Lambda$*. Moreover, if convex and concave conditions are replaced by strictly convex and strictly concave conditions, the saddle point is unique.*

We recast our problem into the saddle point formulation. For this purpose, we set $U = \mathrm{H}_0^1(\Omega) \times \mathrm{L}^2(\Omega)$ and $\Lambda = \mathrm{H}_0^1(\Omega)$ throughout. Let

(2.1)
$$\mathcal{K} = \{(u, g) \in U \mid g \in \mathfrak{U}_{ad} \text{ and } u = u(g) \in \mathrm{H}_0^1(\Omega) \text{ is a solution of (1.6)}\}\,.$$

Now, we introduce a function $\Phi : U \times \Lambda \to \mathbb{R}$ defined by

(2.2) $$\Phi((y, h), q) = a(y, q) - < f, q > -(h, q)\,, \quad \forall ((y, h), q) \in U \times \Lambda\,.$$

Lemma 2.3. *The set* $\mathcal{K}$ *is characterized by*

(2.3) $$(u, g) \in \mathcal{K} \quad \textit{if and only if} \quad \Phi((u, g), q) \le 0\,, \quad \forall q \in \Lambda\,.$$

Proof. Suppose $(u, g) \in \mathcal{K}$. By (1.6), it follows that $\Phi((u, g), q) = 0$ for all $q \in \Lambda$.

Now, assume $\Phi((u, g), q) \le 0$ for all $q \in \Lambda$. Since Φ is 1-homogeneous in second component, we have $\Phi((u, g), -q) = -\Phi((u, g), q) \le 0$. This two inequalities together yield that $\Phi((u, g), q) = 0$ for all $q \in \Lambda$, which implies that

$$a(u, q) = < f, q > +(g, q)\,, \quad \forall q \in \mathrm{H}_0^1(\Omega)\,.$$

This means that $g \in \mathfrak{U}_{ad}$ and $u = u(g)$, so that $(u, g) \in \mathcal{K}$. □

From Lemma 2.3, one can say that

$$\sup_{q \in \Lambda} \Phi((y,h),q) = \begin{cases} 0 & \text{if } (y,h) \in \mathcal{K}\,, \\ +\infty & \text{if } (y,h) \notin \mathcal{K}\,. \end{cases} \tag{2.4}$$

Based on these facts, we introduce the Lagrangian $\mathcal{F}$ over $U \times \Lambda$ associated with the minimization problem by setting up

$$\begin{aligned} \mathcal{F}((y,h),q) &= J(y,h) + \Phi((y,h),q) \\ &= \frac{1}{2}\|y-\omega\|^2 + \frac{\alpha}{2}\|h\|^2 + a(y,q) - < f, q > -(h,q)\,, \end{aligned} \tag{2.5}$$

where $J(y,h)$ is the cost functional in (1.2). Then, from (2.4), our problem (1.9) corresponds to the following minimax problem;

$$\inf_{(y,h)\in\mathcal{K}} J(y,h) = \inf_{(y,h)\in\mathcal{K}} \sup_{q\in\Lambda} \Big(J(y,h) + \Phi((y,h),q) \Big)\,. \tag{2.6}$$

The existence of a saddle point provides a major contribution to seeking of an optimal control.

THEOREM 2.4. *There exists a saddle point* $((u,g),v) \in U \times \Lambda$ *such that*

$$\mathcal{F}((u,g),q) \le \mathcal{F}((u,g),v) \le \mathcal{F}((y,h),v)\,, \quad \forall ((y,h),q) \in U \times \Lambda\,. \tag{2.7}$$

Proof. We shall show that the Lagrangian $\mathcal{F}$ satisfies the conditions of Theorem 2.2. For a sake of brevity, we divide our proof into several steps.

Step I: (Existence of a saddle point in the bounded closed subset) Let ℓ be a suitably chosen positive constant which will be determined later. Let us consider two sets

$$\begin{cases} U_\ell = \{(y,h) \in U \mid \|y\|_1 \le \ell \text{ and } \|h\| \le \ell\}\,, \\ \Lambda_\ell = \{q \in \Lambda \mid \|q\|_1 \le \ell\}\,. \end{cases} \tag{2.8}$$

and let $\mathcal{K}_\ell \equiv \mathcal{K} \cap U_\ell$. We first show that $\mathcal{F}$ restricted to $U_\ell \times \Lambda_\ell$ satisfies all the conditions of Theorem 2.2. Since $\mathcal{F}$ is linear and continuous in the second component, for a given $(y,h) \in U_\ell$, $\mathcal{F}((y,h),\cdot)$ is concave and upper-semicontinuous. Also, since $J(y,h)$ is positive semi-definite and $a(\cdot,\cdot)$ is $\mathrm{H}_0^1(\Omega)$-coercive, obviously $\mathcal{F}(\cdot,q)$ is convex for a given $q \in U_\ell$. Moreover, since $a(\cdot,\cdot)$ is $\mathrm{H}_0^1(\Omega)$-continuous and $J(y,h)$ is continuous at every $(y,h) \in \Lambda_\ell$, $\mathcal{F}(\cdot,q)$ is also continuous. Thus, the conditions [i] and [ii] of Theorem 2.2 are satisfied. For condition [iii], it is sufficient to take $q = 0$ in Λ_ℓ. Finally, if we take any $(u,g) \in \mathcal{K}_\ell \subset U_\ell$, then by (2.4) we have the condition [iv]. Hence from Theorem 2.2 there exists a saddle point $((u_\ell,g_\ell),v_\ell) \in U_\ell \times \Lambda_\ell$ such that

$$(2.9) \quad \begin{cases} J(u_\ell,g_\ell) + \Phi((u_\ell,g_\ell),q) \le J(u_\ell,g_\ell) + \Phi((u_\ell,g_\ell),v_\ell) \\ \qquad\qquad \le J(y,h) + \Phi((y,h),v_\ell)\,, \quad \forall((y,h),q) \in U_\ell \times \Lambda_\ell\,. \end{cases}$$

In the following steps, choosing ℓ large enough we shall show that $(u_\ell,g_\ell) \in U_\ell$ and $v_\ell \in \Lambda_\ell$ are bounded independent of the choice of such an ℓ.

<u>StepII:</u> (Boundedness of g_ℓ)
By the second inequality of (2.9), $\mathcal{F}(\cdot,v_\ell)$ takes a local minimum at $(u_\ell,g_\ell) \in U_\ell$. However, convexity of $\mathcal{F}(\cdot,v_\ell)$ induces a global minimum at (u_ℓ,g_ℓ). Hence, from (2.9) it is followed by

$$(2.10) \quad \begin{cases} J(u_\ell,g_\ell) + \Phi((u_\ell,g_\ell),q) \le J(u_\ell,g_\ell) + \Phi((u_\ell,g_\ell),v_\ell) \\ \qquad\qquad \le J(y,h) + \Phi((y,h),v_\ell)\,, \quad \forall((y,h),q) \in U \times \Lambda_\ell\,. \end{cases}$$

Taking $q = 0$ and $(y,h) = (u,g)$, (2.10) is reduced to

$$J(u_\ell,g_\ell) \le J(u_\ell,g_\ell) + \Phi((u_\ell,g_\ell),v_\ell) \le J(u(g),g) = J(g)\,.$$

Hence, it follows that

$$(2.11) \qquad \Phi((u_\ell,g_\ell),v_\ell) \ge 0 \quad \text{and} \quad J(u_\ell,g_\ell) \le J(g)\,.$$

From the second inequality of (2.11), immediately we have

$$(2.12) \qquad \|g_\ell\| \le \sqrt{\frac{2}{\alpha}J(g)} \equiv c_3\,.$$

Step III: (Boundedness of u_ℓ)
Take $(y,h)=(u,g)\in\mathcal{K}$ and $q=\ell\dfrac{u_\ell}{\|u_\ell\|_1}\in\Lambda_\ell$ in (2.10). Then, since Φ is 1-homogeneous in the second component and $\Phi((u,g),v_\ell)=0$ for $v_\ell\in\Lambda_\ell$, the second inequality of (2.10) is reduced to

$$J(u_\ell,g_\ell)+\frac{\ell}{\|u_\ell\|_1}\Phi((u_\ell,g_\ell),u_\ell)\le J(g)\,. \tag{2.13}$$

Since $J(u_\ell,g_\ell)\ge 0$, from the definition of Φ, (2.13) induces

$$\begin{aligned} a(u_\ell,u_\ell) &\le \frac{\|u_\ell\|_1}{\ell}J(g)+<f,u_\ell>+(g_\ell,u_\ell)\\ &\le \|u_\ell\|_1\Big\{\frac{1}{\ell}J(g)+(\|f\|_{-1}+\|g_\ell\|)\Big\}\,. \end{aligned}$$

Now, applying the coercivity (1.5) of $a(\cdot,\cdot)$ and taking ℓ so that $\ell>1$, this yields that

$$\begin{aligned} \|u_\ell\|_1 &\le {c_2}^{-1}(J(g)+\|f\|_{-1}+\|g_\ell\|)\\ &\le c_2^{-1}(J(g)+\|f\|_{-1}+c_3)\equiv c_4\,. \end{aligned} \tag{2.14}$$

Step IV: (Boundedness of v_ℓ)
For a fixed $v_\ell\in\Lambda_\ell$, let τ be a function on defined on U by

$$\tau(y,h)=\mathcal{F}((y,h),v_\ell)\,.$$

Since (u_ℓ,g_ℓ) is the global minimum of $\tau(\cdot)=\mathcal{F}(\cdot,v_\ell)$, taking a G-derivative of τ at $(u_\ell,g_\ell)\in U_\ell$, we have

$$\tau'((u_\ell,g_\ell);(y,h))=0\,,\quad \forall(y,h)\in U\,.$$

From (2.5), this virtually produces

$$(u_\ell-\omega,y)+(\alpha g_\ell,h)+a(y,v_\ell)-(h,v_\ell)=0\,,\quad \forall(y,h)\in U\,.$$

If we plug $(y,h)=(v_\ell,g_\ell)$, this is reduced to

$$a(v_\ell,v_\ell)+\alpha\|g_\ell\|^2=(g_\ell,v_\ell)-(u_\ell-\omega,v_\ell)\,.$$

Finally, since $\alpha > 0$, applying the coercivity (1.5) of $a(\cdot,\cdot)$ we obtain

$$\begin{aligned} c_2\|v_\ell\|_1^2 &\le a(v_\ell, v_\ell) + \alpha\|g_\ell\|^2 \\ &\le (\|g_\ell\| + \|u_\ell - \omega\|)\|v_\ell\|_1 \,. \end{aligned}$$

From this, it is followed by

$$\begin{aligned} \|v_\ell\|_1 &\le {c_2}^{-1}(\|g_\ell\| + \|u_\ell - \omega\|) \\ &\le {c_2}^{-1}(c_3 + c_4 + \|\omega\|) \equiv c_5 \,. \end{aligned} \tag{2.15}$$

Step V: (Choice of ℓ and $u_\ell = u(g_\ell)$)
Let us consider U_ℓ and Λ_ℓ with $\ell > \max\{1, c_3, c_4, 2c_5\}$. Verifying $u_\ell = u(g_\ell)$ is equivalent to showing that

$$\phi((u_\ell, g_\ell), q) = 0\,, \quad \forall q \in \Lambda\,. \tag{2.16}$$

From (2.11), we already have $\Phi((u_\ell, g_\ell), v_\ell) \ge 0$. Since $q = 2v_\ell \in \Lambda$ satisfies $\|q\|_1 = 2\|v_\ell\|_1 \le 2c_5 < \ell$, $2v_\ell$ belongs to Λ_ℓ. Now, plugging $q = 2v_\ell$ into the first inequality of (2.9), we attain

$$\Phi((u_\ell, g_\ell), v_\ell) \le 0\,. \tag{2.17}$$

Recoursing once again to the first inequality of (2.9), we derive that

$$\Phi((u_\ell, g_\ell), q) = 0\,, \quad \forall q \in \Lambda_\ell\,. \tag{2.18}$$

If $q \in \Lambda$ does not belongs to Λ_ℓ, $\ell\dfrac{q}{\|q\|_1} \in \Lambda_\ell$ will take place the role due to the homogeneity of Φ in the second component.

Finally, since $u_\ell = u(g_\ell)$, from (2.9) of Step I one can conclude the existence of a saddle point of the Lagrangian $\mathcal{F}$. □

Previous theorem states that a saddle point $((u,g),v) \in U \times \Lambda$ and that $u = u(g)$, i.e., $(u,g) \in \mathcal{K}$ is the solution of the state equation (1.6). The second component v corresponding to *a Lagrange multiplier* of the Lagrangian $\mathcal{F}$ is a dual variable of the problem.

2.2. The Optimality System

Let us concentrate on (2.7). To obtain the characterization for the control parameter, we need to introduce a function $\widehat{\mathcal{F}} : V \times \mathrm{L}^2(\Omega) \times V \to \mathbb{R}$ which is induced from $\mathcal{F}$ by

$$\widehat{\mathcal{F}}(y,h,q) = \mathcal{F}((y,h),q) \quad \text{for } (y,h,q) \in V \times \mathrm{L}^2(\Omega) \times V\,. \tag{2.19}$$

This is indispensable to achieve governing equations for each separate variables. The second inequality of (2.7) tells us that if we set $h = g \in \mathfrak{U}_{ad}$, $y = u$ is the minimizer of a function $\zeta : V \to \mathbb{R}$ defined by

$$\begin{aligned} \zeta(y) &= \widehat{\mathcal{F}}(y,g,v) \\ &= \frac{1}{2}\|y-\omega\|^2 + \frac{\alpha}{2}\|g\|^2 + a(y,v) - < f, v > -(g,v)\,. \end{aligned} \tag{2.20}$$

So, taking a G-derivative of ζ at u, we derive that

$$\zeta'(u;\phi) = (u-\omega,\phi) + a(\phi,v) = 0\,, \quad \forall \phi \in V\,.$$

So, the dual variable $v \in V$ should satisfy the following *adjoint equation*

$$a(\phi,v) = -(u-\omega,\phi)\,, \quad \forall \phi \in V\,. \tag{2.21}$$

Since $a(\cdot,\cdot)$ is V-continuous and coercive, for a given $u \in V$, there exists a unique solution $v \in V$ of the adjoint equation.

Now, let us derive a necessary and sufficient condition for the optimal solution g of our problem. We once again recourse to the second inequality of (2.7). If we take $y = u$ in the second inequality, $h = g \in \mathfrak{U}_{ad}$ is the minimizer of a function $\eta : \mathrm{L}^2(\Omega) \to \mathbb{R}$ defined by

$$\begin{aligned} \eta(h) &= \widehat{\mathcal{F}}(u,h,v) \\ &= \frac{1}{2}\|u-\omega\|^2 + \frac{\alpha}{2}\|h\|^2 + a(u,v) - < f, v > -(h,v)\,. \end{aligned} \tag{2.22}$$

The necessary and sufficient condition is derived from Lemma 2.1; g is the optimal solution if and only if

$$g \in \mathfrak{U}_{ad} \text{ and } \eta'(g;h-g) = (\alpha g - v, h-g) \geq 0\,, \quad \forall h \in \mathfrak{U}_{ad}\,.$$

Let us summarize this results in the following.

THEOREM 2.5. (**The optimality system**) The optimal solution $g \in \mathfrak{U}_{ad}$ of the problem (1.9) should satisfy the inequality

$$g \in \mathfrak{U}_{ad}; \quad (\alpha g - v, h - g) \geq 0\,, \quad \forall h \in \mathfrak{U}_{ad}\,, \tag{2.23}$$

where $u = u(g)$ is the solution of the *state equation*

$$u \in V; \quad a(u, \psi) = < f, \psi > + (g, \psi)\,, \quad \forall \psi \in V \tag{2.24}$$

and v is the solution of the *adjoint equation*

$$v \in V; \quad a(\phi, v) = -(u - \omega, \phi)\,, \quad \forall \phi \in V\,. \tag{2.25}$$

The system (2.23)–(2.25) is called the *optimality system* to the problem. We shall employ this optimality system to find an efficient scheme to the approximation in the next section.

3. Approximation

We do make use of the optimality system (2.23)–(2.25) to set up an algorithm to approximate an optimal solution. One may establish several different algorithms based on suitable optimization tools such as Uzawa's method, projected gradient method or contraction mapping method, etc., (see [2] or [5]). Since we have the inequality condition (2.23) for the optimality, we need to recharacterize the optimality by utilizing the state equation (2.24) and adjoint equation (2.25). In the first place, we remove the state variable by employing the dual variable in the operational approach to the optimality system and then we derive a pure optimization problem only the control parameter is concerned with. The gradient of J is ensued as a natural result, from it we propose a projected gradient method to attain an optimal solution.

3.1. Operational Approach to the Optimality System

Let $\mathcal{A}$ be a Laplace operator $-\Delta$ in Ω with homogeneous Dirichlet boundary condition imposed. Then, $\mathcal{A}$ is an operator associated with the bilinear form $a(\cdot, \cdot)$ through

$$a(u, v) = < \mathcal{A}u, v >, \quad \forall u, v \in V\,. \tag{3.1}$$

Since $a(\cdot,\cdot)$ is V-continuous and V-coercive, $\mathcal{A}$ is a bijective continuous operator from V to V^* with a continuous inverse $\mathcal{A}^{-1}$ with the operator norm $\|\mathcal{A}^{-1}\| \leq \dfrac{1}{c_2}$. By using Lemma 1.2, for a given $g \in \mathfrak{U}_{ad}$ one can write

$$u(g) = \mathcal{A}^{-1}(f+g) \tag{3.2}$$

whence the functional $J(g)$ of (1.8) can be rewritten by

(3.3)

$$\begin{aligned} J(g) &= \frac{1}{2}\|\mathcal{A}^{-1}(f+g)-\omega\|^2 + \frac{\alpha}{2}\|g\|^2 \\ &= \frac{1}{2}\|\mathcal{A}^{-1}(g)\|^2 + \frac{\alpha}{2}\|g\|^2 + (\mathcal{A}^{-1}f-\omega, \mathcal{A}^{-1}g) + \frac{1}{2}\|\mathcal{A}^{-1}f-\omega\|^2 \,. \end{aligned}$$

Having expressed J as the sum of quadratic forms, we can easily compute the G-derivative of J as follows.

$$\begin{aligned} J'(g;h) &= (\mathcal{A}^{-1}g, \mathcal{A}^{-1}h) + (\mathcal{A}^{-1}f-\omega, \mathcal{A}^{-1}h) + (\alpha g, h) \\ &= (\mathcal{A}^{-1}(f+g)-\omega, \mathcal{A}^{-1}h) + (\alpha g, h) \\ &= (u(g)-\omega, \mathcal{A}^{-1}h) + (\alpha g, h)\,, \quad \forall h \in \mathrm{L}^2(\Omega)\,. \end{aligned} \tag{3.4}$$

Hence, one can write the gradient of J as

$$\nabla J(g) = \mathcal{A}^{-1^*}(u(g)-\omega) + \alpha g\,, \tag{3.5}$$

where $\mathcal{A}^{-1^*}$ denotes the adjoint of $\mathcal{A}^{-1}$.

THEOREM 3.1. *There exists a unique optimal control g^*.*

Proof. From (3.3) and (3.4), we have

$$\begin{aligned} <\nabla J(g) - \nabla J(h), g-h> &= J'(g; g-h) - J'(h; g-h) \\ &= \{(u(g)-\omega, \mathcal{A}^{-1}(g-h)) + (\alpha g, g-h)\} \\ &\qquad - \{(u(h)-\omega, \mathcal{A}^{-1}(g-h)) + (\alpha h, g-h)\} \\ &= (u(g)-u(h), \mathcal{A}^{-1}(g-h)) + (\alpha(g-h), g-h) \\ &= (u(g)-u(h), \mathcal{A}^{-1}(f+g) - \mathcal{A}^{-1}(f+h)) + (\alpha(g-h), g-h) \\ &= (u(g)-u(h), u(g)-u(h)) + (\alpha(g-h), g-h) \\ &= \|u(g)-u(h)\|^2 + \alpha\|g-h\|^2 \end{aligned}$$

for all $g, h \in \mathfrak{U}_{ad}$. Hence, ∇J is a strictly monotone operator. Now, for every $g \in \mathfrak{U}_{ad}$, since $\mathcal{A}^{-1} \in \mathcal{L}(V^*, V)$, there exists a positive constant $c(= \dfrac{1}{c_2})$ such that

$$|(u(g) - \omega, \mathcal{A}^{-1}g)| \leq c\|u(g) - \omega\|\|g\| .$$

Hence, it follows that

$$\begin{aligned} < \nabla J(g), g > &= J'(g; g) \\ &= (u(g) - \omega, \mathcal{A}^{-1}g) + (\alpha g, g) \\ &\geq \alpha\|g\|^2 - c\|u(g) - \omega\|\|g\| . \end{aligned}$$

Since α is a positive constant, it is obvious that ∇J is coercive. Since $\mathfrak{U}_{ad}$ is a closed convex set, we have a unique optimal control by Lemma 2.1. □

The optimal control $g^* \in \mathfrak{U}_{ad}$ is characterized by the variational inequality

$$J'(g^*; g - g^*) \geq 0 , \quad \forall g \in \mathfrak{U}_{ad} . \tag{3.6}$$

However, this inequality is very difficult to interpret in itself. In control sense, this difficulty comes from the disparity to compare observation data with elements of the control space. One can derive an equivalent characterization on the control space $\mathfrak{U}_{ad}$ by employing adjoint variable. Let $\mathcal{A}^* \in \mathcal{L}(V, V^*)$ be an adjoint operator of $\mathcal{A}$ given by

$$< \mathcal{A}u, v >=< u, \mathcal{A}^*v > , \quad \forall u, v \in V .$$

Since $\mathcal{A}$ is an operator associated with the bilinear form $a(\cdot, \cdot)$, $\mathcal{A}^*$ is an operator adjoint to $a(\cdot, \cdot)$, i.e.,

$$a(u, v) =< u, \mathcal{A}^*v > , \quad \forall u, v \in V .$$

Hence, $\mathcal{A}^*$ also corresponds to a Laplace operator $-\Delta$ with homogeneous Dirichlet boundary condition assigned. In our case, virtually

$\mathcal{A}^*$ coincides with $\mathcal{A}$. To induce a more convenient form to characterize the optimal control, from (2.25) we introduce an adjoint state $v = v(g) \in V$ as the unique solution of

$$\mathcal{A}^* v = -(u(g) - \omega)\,. \tag{3.7}$$

Then, from (3.4) or (3.5), we obtain the condition for the optimal solution g^*

$$\begin{aligned} J'(g^*; g - g^*) &= (u(g^*) - \omega, \mathcal{A}^{-1}(g - g^*)) + (\alpha g^*, g - g^*) \\ &=< -\mathcal{A}^* v(g^*), \mathcal{A}^{-1}(g - g^*) > +(\alpha g^*, g - g^*) \\ &=< -v(g^*), \mathcal{A}\mathcal{A}^{-1}(g - g^*) > +(\alpha g^*, g - g^*) \\ &= (\alpha g^* - v(g^*), g - g^*) \\ &\geq 0\,, \quad \forall g \in \mathfrak{U}_{ad}\,. \end{aligned} \tag{3.8}$$

This is nothing but an optimality condition (2.23) and from this the gradient of J is simply represented by

$$\nabla J(g) = \alpha g - v(g)\,, \tag{3.9}$$

where $v = v(g)$ is the solution of the adjoint equation (3.7). This information can be directly employed in establishing a suitable algorithm to achieve an optimal control. We shall propose an algorithm based on the projected gradient method in the next section.

Let us summarize above results as follows.

Theorem 3.2. (**Optimality system revisited**) *Let the optimal control problem be given in (1.8). Then, a necessary and sufficient condition for $g^* \in \mathfrak{U}_{ad}$ to be an optimal control is that it satisfies the following system*

$$\begin{cases} u(g^*) \in V, \quad \mathcal{A}u(g^*) = f + g^*\,, \\ v(g^*) \in V, \quad \mathcal{A}^* v(g^*) = -(u(g^*) - \omega)\,, \\ (\alpha g^* - v(g^*), g - g^*) \geq 0\,, \quad \forall g \in \mathfrak{U}_{ad} \end{cases} \tag{3.10}$$

Remark. When $\mathfrak{U}_{ad} = \mathrm{L}^2(\Omega)$, the optimal solution is given by $g^* = \dfrac{1}{\alpha} v(g^*)$, where $v = v(g^*)$ is a unique solution of

$$\begin{cases} -\Delta v = -(u(g^*) - \omega) \quad \text{in } \Omega\,, \\ \qquad v = 0 \quad \text{on } \partial\Omega\,. \end{cases}$$

Here, $u = u(g^*)$ is a unique solution of (1.6) with g replaced by g^*. Furthermore, since $\mathcal{A}^*$ is well-posed, for each $g \in \mathfrak{U}_{ad}$ there exists a positive constant κ_2 such that

$$\|v(g)\|_2 \leq \kappa_2 \|u(g) - \omega\|\,. \tag{3.11}$$

Remark. The condition $\alpha > 0$ has been essential in verifying that the gradient of J is coercive. If $\alpha = 0$, one can show that an optimal control may not exist. For a special case of this, one may refer to [8].

3.2. Algorithm based on the Projected Gradient Method

Since we have the gradient of J close at hand in (3.9), one can apply the projected gradient method to get an approximation to the optimal solution g^*. Let $\mathcal{P} : L^2(\Omega) \to \mathfrak{U}_{ad}$ be a projection operator so that $\|\mathcal{P}\| \leq 1$.

Algorithm. *For a suitable fixed step size $\rho > 0$, we set the descent direction of J at $g_m \in \mathfrak{U}_{ad}$ by*

$$w_m = \frac{\alpha g_m - v(g_m)}{\|\alpha g_m - v(g_m)\|}\,.$$

For an arbitrary chosen starting point $g_0 \in \mathfrak{U}_{ad}$, evaluate the iterates

$$g_{m+1} = \mathcal{P}(g_m - \rho w_m) \quad \textit{for } m = 0, 1, 2, \ldots, \tag{3.12}$$

until the stopping criterion is met.

We show that this algorithm yields a convergent sequence to the optimal solution. As a preliminary, we need the following estimates.

LEMMA 3.3. *There exists a positive constant* κ_3 *such that*

$$\|\nabla J(g) - \nabla J(h)\| \le \kappa_3 \|g-h\|\,, \quad \forall g, h \in \mathfrak{U}_{ad}\,. \tag{3.13}$$

Proof. Using (3.9) and (3.11), for each g, $h \in \mathfrak{U}_{ad}$ it follows that

$$\begin{aligned}\|\nabla J(g) - \nabla J(h)\| &= \|(\alpha g - v(g)) - (\alpha h - v(h))\| \\ &\le \alpha\|g-h\| + \|v(g)-v(h)\| \\ &\le \alpha\|g-h\| + \|v(g)-v(h)\|_2 \\ &\le (\alpha+\kappa_2)\|g-h\|\,.\end{aligned}$$

So, $\kappa_3 = \alpha + \kappa_2$ will carry out the proof. □

THEOREM 3.4. *If* $0 < \rho < \dfrac{2\alpha}{(\kappa_3)^2}$, *then the proposed algorithm ensures the convergence to the optimal solution.*

Proof. For simplicity, we may assume the normalized gradient so that $\|\nabla J(g)\| = 1$ for each $g \in \mathfrak{U}_{ad}$. For a fixed $0 < \rho < \dfrac{2\alpha}{(\kappa_3)^2}$, let us consider the mapping $\mathfrak{T}_\rho : \mathfrak{U}_{ad} \to \mathfrak{U}_{ad}$ which is defined by

$$\mathfrak{T}_\rho(g) = \mathcal{P}(g - \rho\nabla J(g))\,. \tag{3.14}$$

If we show that $\mathfrak{T}_\rho$ is a contraction mapping, then algorithm (3.12) yields a convergent sequence to a unique fixed point g^* of $\mathfrak{T}_\rho$. So, g^* satisfies

$$g^* = \mathcal{P}(g^* - \rho\nabla J(g^*))\,.$$

Then, since $\mathfrak{T}_\rho$ is a projection onto a closed convex subset, this yields the variational inequality (3.6). So, such a g^* is an optimal solution.

Now, it remains to verify that $\mathfrak{T}_\rho$ is a contraction. Since $\|\mathcal{P}\| \le 1$, using Lemma 3.3 it follows that

$$\begin{aligned}\|\mathfrak{T}_\rho(g) - \mathfrak{T}_\rho(h)\|^2 &= \|\mathcal{P}(g-\rho\nabla J(g)) - \mathcal{P}(h-\rho\nabla J(h))\|^2 \\ &\le \|(g-h) - \rho(\nabla J(g) - \nabla J(h))\|^2 \\ &= \|g-h\|^2 - 2\rho(\nabla J(g) - \nabla J(h), g-h) \\ &\qquad\qquad + \rho^2\|\nabla J(g) - \nabla J(h)\|^2 \\ &\le \Big\{1 - 2\alpha\rho + (\kappa_3)^2\rho^2\Big\}\|g-h\|^2\,.\end{aligned}$$

Here, we also have used the inequality $(\nabla J(g) - \nabla J(h), g-h) \geq \alpha\|g-h\|^2$ which was derived in the proof after Theorem 3.1. So, selecting

$$0 < \rho < \frac{2\alpha}{(\kappa_3)^2}\,, \tag{3.15}$$

$\mathfrak{T}_\rho$ is a contraction. □

We note that each iterate of (3.12) requires solutions $u(g_m)$ and $v(g_m)$ of the state equation and the adjoint equation, respectively. Hence, the computations may be somewhat costly.

References

1. A. Bermudez and C. Saguez, *Optimality conditions for optimal control problems of variational inequalities*, Control and Information Sciences, vol. 97, 1986, pp. 143–153.
2. J. Cea, *Lectures on optimization theory and algorithm*, Springer-Verlag, New York, 1978.
3. P.G. Ciarlet, *The finite element methods for elliptic problems*, North–Holland, Amsterdam, 1980.
4. R. Dautray and J.L. Lions, *Mathematical analysis and numerical methods for science and technology: Vol2; Functional and variational methods*, Springer–Verlag, Berlin, 1988.
5. I. Ekeland and R. Temam, *Convex analysis and variational problems*, North–Holland, Oxford, 1976.
6. J.L. Lions, *Some aspects of the optimal control of distributed parameter systems*, Siam, Philadelphia, 1972.
7. M. Renardy and R.C. Rogers, *Introduction to partial differential equations*, Springer–Verlag, New York, 1993.
8. R.E. Showalter, *Hilbert space methods for partial differential equations*, Austin Texas, 1994, Preprint from Lecture Notes.
9. F. Treves, *Basic linear partial differential equations*, North–Holland, Amsterdam, 1975.

Hongchul Kim
Department of Mathematics
Kangnŭng National University
Kangnŭng, Korea 210–702
E-mail: hongchul@knusun.kangnung.ac.kr

Gye-Soo Rim

Department of Mathematics
Kangnŭng National University
Kangnŭng, Korea 210–702
E-mail: rgyesoo@knusun.kangnung.ac.kr