Appl. Math. and Comp. Sci., 1996, vol.6, No.4, 753–758

AN EXAMPLE OF NON-EXISTENCE OF A CONE APPROXIMATION TO THE SET OF FEASIBLE STATES FOR AN OPTIMAL CONTROL PROBLEM†

ULDIS RAITUMS*

The paper considers optimal control problems for linear elliptic systems with a standard set S of admissible controls σ. In the case where the principal part of the operator depends on controls, it is shown that the set $Z(S)$ of all solutions $u(\sigma)$ of the state equations with $\sigma \in S$ cannot be approximated, in general, by cones, i.e. for a given $\sigma_0 \in S$ there is, in general, neither element h nor family $\{\sigma_\epsilon\} \subset S$ such that $u(\sigma_\epsilon) = u(\sigma_0) + \epsilon h + o(\epsilon)$ as $\epsilon \to 0$.

1. Introduction

In a major part of investigations concerning the necessary optymality conditions as well as in the sensitivity analysis for optimal control problems an assumption has been made on the families $\{\sigma_\epsilon\}$ of admissible controls that the corresponding states $u = u(\sigma)$ possess the property

$$u(\sigma_\epsilon) = u(\sigma_0) + \epsilon h + o(\epsilon) \tag{1}$$

as $\epsilon \to 0$, see e.g. (Warga, 1972) for the case of ordinary differential equations. The same scheme has been employed for partial differential equations (Lions, 1968; Raitums, 1989).

As it appears now (Raitums, 1994) many results of this kind were due to the fact that the validity of the necessary optimality condition in the form of the Pontryagin Maximum Principle (or analogues of this Principle) implies the relaxability via convexification of optimal control problems under consideration. In turn, property (1) is a direct result of the convexity of the set of admissible operators and of the Implicit-Function Theorem.

On the other hand, until now there has not been any clear idea of the type of necessary optimality conditions for optimal control problems governed by systems of partial differential equations with the main part of the differential operator depending on controls from a nonconvex set.

† This research was supported by the International Science Foundation (grant No. LF6000) and by the Latvian Council of Sciences (grant No. 93.604)

* Institute of Mathematics and Computer Science, University of Latvia, LV-1459 Riga, Latvia, e-mail: raitums@cclu.lv

In this paper, we give a simple example of an elliptic system of equations

$$\operatorname{div} A(\sigma)\nabla \boldsymbol{u} = \operatorname{div} \boldsymbol{f}, \quad x \in \Omega \subset \mathbb{R}^n, \quad \boldsymbol{u} = (u_1, \dots, u_n) \in \left[H_0^1(\Omega)\right]^n \tag{2}$$

with a standard set of admissible controls

$$S := \left\{\sigma \in L_2(\Omega) \mid \sigma(x) = 0 \text{ or } 1, \ x \in \Omega\right\}$$

such that for some $\sigma_0 \in S$ there does not exist a family $\{\sigma_\epsilon\} \subset S$ such that property (1) holds with a non-trivial element h.

Roughly speaking, our example shows that the set $Z(S)$ of all solutions to (2) with $\sigma \in S$ has some properties which are similar to those of the set S itself. We believe that this fact will be helpful to explain the difficulties which arise in the investigations of optimal control problems for systems of partial differential equations.

2. Notation and Preliminaries

Let Ω be a bounded domain in $\mathbb{R}^n$ with a Lipschitz boundary $\partial\Omega$ and let D be a strictly interior subdomain, $\overline{D} \subset \Omega$. We introduce the following spaces:

- $\boldsymbol{L} := \left[L_2(\Omega)\right]^{n \times n}$ is the space of $n \times n$ matrix-functions $\boldsymbol{f} = (f_j^i)$, $i,j = 1, \dots, n$, with square-integrable elements. We will denote by f^i, $i = 1, \dots, n$ the rows of a matrix-function $\boldsymbol{f}$.
- $\boldsymbol{H} := \left[H_0^1(\Omega)\right]^n$ is the Sobolev space of vector-functions $\boldsymbol{u} = (u^1, \dots, u^n)$ whose components belong to $H_0^1(\Omega)$.
- $\boldsymbol{G} := \left\{\boldsymbol{f} \in L \mid \boldsymbol{f} = \operatorname{grad} \boldsymbol{u} \text{ for some } \boldsymbol{u} \in H\right\}$.

We will denote by P the operator of the orthogonal projection of $\boldsymbol{L}$ onto $\boldsymbol{G}$. The notation $|Q|$ means the Lebesgue measure of a set Q.

In the sequel, we shall need the following result:

Lemma 1. (Nečas, 1965) *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^n$. Then there exists a constant $c(\Omega)$ depending only on Ω such that for every $a \in L_2(\Omega)$ with*

$$\int_\Omega a \, \mathrm{d}x = 0$$

we have

$$\|a\|_{L_2(\Omega)} \le c(\Omega) \sup_{\boldsymbol{u} \in \boldsymbol{H}, \, \|\boldsymbol{u}\| \le 1} \sum_{i=1}^n \int_\Omega a u_{x_i}^i \, \mathrm{d}x \tag{3}$$

Since P is the operator of the orthogonal projection of $\boldsymbol{L}$ on $\boldsymbol{G}$, the elliptic system (the number of equations is equal to the dimension of the space)

$$\operatorname{div} A\nabla \boldsymbol{u} = \operatorname{div} \boldsymbol{f}, \qquad x \in \Omega \tag{4}$$

with respect to $\boldsymbol{u} \in \boldsymbol{H}$ is equivalent to the equation

$$P\Big[A\nabla \boldsymbol{u} - \boldsymbol{f}\Big] = 0 \tag{5}$$

In what follows, we will often use the form of eqn. (5) instead of (4).

Let D be a strictly interior subdomain of Ω and

$$S := \Big\{\sigma \in L_2(\Omega) \ \big|\ \sigma(x) = 0 \text{ or } 1 \text{ if } x \in D,\ \sigma(x) = 0 \text{ if } x \in \Omega \setminus D\Big\}$$

denote the set of admissible controls. Let $\varsigma \in C_0^\infty(\Omega)$ be a function such that $\varsigma(x) = 1$ if $x \in D$, and $\boldsymbol{f}_0 = (f_0^1, \dots, f_0^n)$, $f_0^i(x) = \operatorname{grad}\big(x_i\varsigma(x)\big)$, $i = 1, \dots, n$.

Consider the family of elliptic systems

$$\operatorname{div}\Big(1 + \sigma(x)\Big)\nabla \boldsymbol{u} = \operatorname{div} \boldsymbol{f}_0, \quad x \in \Omega, \quad \boldsymbol{u} \in \boldsymbol{H} \tag{6}$$

or, in our notation, the family of equations

$$P\Big[(1 + \sigma)\nabla \boldsymbol{u} - \boldsymbol{f}_0\Big] = 0 \tag{7}$$

depending on $\sigma \in S$. The solution of eqn. (7) corresponding to a chosen $\sigma \in S$ will be denoted by $\boldsymbol{u}(\sigma)$.

Our basic result is as follows:

Proposition 1. *For the family of equations (7) and for $\sigma_0 = 0$ neither non-trivial element $\boldsymbol{h} \in \boldsymbol{H}$ nor family $\{\sigma_\epsilon\} \subset S$ exists such that*

$$\boldsymbol{u}(\sigma_\epsilon) = \boldsymbol{u}(\sigma_0) + \epsilon \boldsymbol{h} + o(\epsilon)$$

where

$$\frac{\|o(\epsilon)\|_{\boldsymbol{H}}}{\epsilon} \to 0 \quad \textit{as} \quad \epsilon \to 0$$

3. Proof of the Basic Result

We begin by modifying the result of Lemma 1.

Lemma 2. *Let E be the identity matrix in $\mathbb{R}^{n\times n}$ and let $\{\sigma_k\} \subset S$ be a sequence such that*

$$P\sigma_k E \to 0 \quad \textit{in} \quad \boldsymbol{L} \quad \textit{as} \quad k \to \infty$$

Then $\sigma_k \to 0$ in $L_2(\Omega)$ as $k \to \infty$.

Proof. Lemma 1 implies that for every $a \in L_2(\Omega)$ with

$$\int_\Omega a \,\mathrm{d}x = 0$$

we have

$$\|a\|_{L_2(\Omega)} \le c_1 \|PaE\|_{\boldsymbol{L}}$$

where the constant c_1 depends only on $c(\Omega)$ and n.

Since for every constant $c \in \mathbb{R}$

$$PcE = 0$$

we conclude that

$$\left\| \sigma_k - \frac{1}{|\Omega|} \int_\Omega \sigma_k \, \mathrm{d}x \right\|_{L_2(\Omega)} \le c_1 \left\| P \left[\left(\sigma_k - \frac{1}{|\Omega|} \int_\Omega \sigma_k \, \mathrm{d}x \right) E \right] \right\|_{\boldsymbol{L}}$$

$$= c_1 \|P\sigma_k E\|_{\boldsymbol{L}} \to 0 \quad \text{as} \quad k \to \infty$$

From this and from the definition of the set S (the functions $\sigma_k \in S$ are equal to zero outside the domain D) the statement of the lemma follows. ■

Let us suppose that for $\boldsymbol{u}_0 = \boldsymbol{u}(\sigma_0)$ (with $\sigma_0 = 0$) there exist a nonzero element $\boldsymbol{h} \in \boldsymbol{H}$ and a family $\{\sigma_t\} \subset S$, $0 < t < t_0$, such that

$$\boldsymbol{u}(\sigma_t) = \boldsymbol{u}_0 + t\boldsymbol{h} + o(t)$$

From the relationships

$$P\Big[(1+\sigma_t)\nabla \boldsymbol{u}(\sigma_t) - \boldsymbol{f}_0 - 0\Big] = 0$$

and

$$P[\nabla \boldsymbol{u}_0 - \boldsymbol{f}_0] = 0$$

it follows that

$$P[\sigma_t \nabla \boldsymbol{u}_0] + tP[\sigma_t \nabla \boldsymbol{h}] + tP\nabla \boldsymbol{h} + P\Big[(1+\sigma_t)o(t)\Big] = 0 \tag{8}$$

Since P is bounded, we see that

$$P[\sigma_t \nabla \boldsymbol{u}_0] \to 0 \quad \text{as} \quad t \to 0$$

This, together with Lemma 2 and the fact that $\nabla \boldsymbol{u}_0(x) = E$ in D, implies

$$\|\sigma_t\|_{L_2(\Omega)} \to 0 \quad \text{as} \quad t \to 0$$

On the other hand, from the convergence $\sigma_t \to 0$ as $t \to 0$ it follows that

$$P[\sigma_t \nabla \boldsymbol{h}] \to 0 \quad \text{as} \quad t \to 0$$

Thus, from (8) we obtain

$$P[\sigma_t \nabla \boldsymbol{u}_0] + t\,P\,\nabla \boldsymbol{h} = o(t) \tag{9}$$

where $\|o(t)\|_{\boldsymbol{H}}/t \to 0$ as $t \to 0$.

The element $P\nabla \boldsymbol{h}$ is equal to $\nabla \boldsymbol{h}$ and, therefore, is not equal to zero. The functions σ_t are equal to zero outside the set D, but in D we have $\nabla \boldsymbol{u}_0(x) = E$ if $x \in D$. Hence, the relationship (9) gives

$$P[\sigma_t E] = -t\nabla \boldsymbol{h} + o(t) \tag{10}$$

Set

$$a_t := \sigma_t - \frac{1}{|\Omega|}\int \sigma_t \,\mathrm{d}x$$

Then from (10) and Lemma 2 we have

$$t\|\boldsymbol{h}\|_{\boldsymbol{H}} - o(t) \leq \|a_t\|_{L_2(\Omega)} \leq c_1 t + o(t)$$

This means that for some subsequence $\{a_{t_k}\} \subset \{a_t\}$ we obtain

$$\|a_{t_k}\|_{L_2(\Omega)} = dt_k + o(t_k), \quad k = 1, 2, \ldots \tag{11}$$

with some constant $d > 0$.

Hence, by virtue of (10) and (11),

$$\nabla \boldsymbol{h} = -P\left[\frac{1}{t_k} a_{t_k} E\right] + \gamma(t_k) \tag{12}$$

where $\gamma(t_k) \to 0$ as $t_k \to 0$.

The sequence $\{a_{t_k}/t_k\}$ is bounded in $L_2(\Omega)$ and, without loss of generality, we can assume that this sequence converges weakly in $L_2(\Omega)$ to some element $b_0 \in L_2(\Omega)$.

Since σ_t are the characteristic functions of subsets of D and $\sigma_t \to 0$ strongly in $L_2(\Omega)$ as $t \to 0$, simple calculations give

$$a_{t_k} = \chi_\epsilon - \frac{\epsilon}{|\Omega|}$$

where χ_ϵ's are the characteristic functions of sets with measure equal to ϵ and

$$\epsilon = d^2 {t_k}^2 + o\Big(|t_k|^2\Big)$$

Hence

$$\frac{a_{t_k}}{t_k} \to 0 \quad \text{weakly in} \quad L_2(\Omega) \quad \text{as} \quad t_k \to 0$$

i.e. $b_0 = 0$.

The operator P is weakly continuous. Therefore, by passing to the limit $t_k \to 0$ in the relationship (12) we obtain

$$\nabla \boldsymbol{h} = 0$$

that contradicts the assumption that $\boldsymbol{h} \neq 0$. ∎

4. One-Dimensional Case

The statement of Proposition 1 can be, in some sense, illustrated by the following one-dimensional example.

Consider the family of two-point boundary-value problems

$$\begin{aligned} &(\sigma y')' = f', \quad 0 < x < 1 \\ &y(0) = y(1) = 0 \end{aligned} \tag{13}$$

where, by definition,

$$y' = \frac{\partial}{\partial x} y$$

and

$$\sigma \in S_1 := \Big\{\sigma \in L_2(0,1) \mid \sigma(x) = \sigma_- \quad \text{or} \quad \sigma_+\Big\}$$

with $0 < \sigma_- < \sigma_+$.

For the solutions of these boundary-value problems we have the explicit formulae

$$y'(\sigma) = \frac{1}{\sigma}\Big[f - c(\sigma)\Big], \qquad c(\sigma) = \int_0^1 \frac{f}{\sigma}\,\mathrm{d}x \left(\int_0^1 \frac{1}{\sigma}\,\mathrm{d}x\right)^{-1}$$

If the function f is not equal to a constant in sets with nonzero measure, then the set $\{y'(\sigma) \in L_2(0,1) \mid \sigma \in S_1\}$ has the structure which is very similar to the structure of the set S_1 itself. It is clear that for a fixed $\sigma_0 \in S_1$, there are no families $\{\sigma_t\} \subset S_1$ such that

$$\sigma_t = \sigma_0 + ta + o(t)$$

with some nonzero element $a \in L_2(0,1)$.

References

Lions J.-L. (1968): *Contrôle Optimal des systèmes gouvernés par des équations aux derivées partielles.* — Paris: Dunod.

Nečas J. (1965): *Equations aux derivées partielles.* — Presses de l'Université de Montréal.

Raitums U. (1989): *Optimal Control Problems for Elliptic Equations.* — Riga: Zinatne, (in Russian).

Raitums U. (1994): *The maximum principle and the convexification of optimal control problems.* — Control and Cybernetics, v.23, No.4, pp.746–760.

Warga J. (1972): *Optimal Control of Differential and Functional Equations.* — New York: Academic Press.

Received: December 5, 1995
Revised: May 20, 1996