Dynamics and Control 2, 331–348 (1992)

A Computational Procedure for Suboptimal Robust Controls*

V. REHBOCK, K.L. TEO, AND L.S. JENNINGS
Department of Mathematics, University of Western Australia, Nedlands, WA, 6009 Australia

Received June 10, 1991; revised November 18, 1991
Editor: N.U. Ahmed

Abstract. We consider a class of optimal control problems that depend on a set of scalar parameters which could have some uncertainty as to their exact values. We show how to compute the control functions given that we wish to balance two objectives. The first is the original objective while the second is the variation of the original objective with respect to the scalar parameters. That is we wish to move the controls to a position where there is less variation with respect to uncertainty in the scalar parameters, perhaps at the expense of the original objective. The gradient of the combined objective is derived and the method demonstrated using two examples.

1. Introduction

Consider a class optimal control problems involving uncertain coefficients. For specified values of these coefficients, let u^* be an optimal control and J^* the corresponding value of the objective functional. We refer to this u^* as the base control. In this article, we wish to consider the question regarding the sensitivity of J^* with respect to these uncertain coefficients. More precisely, if we use the same base control u^*, but make a small change in the values of the coefficients, how will the value of J^* change in response? Clearly, it is desirable that this change be small. The aim of this article is to propose a computational approach for solving this class of optimal control problems in such a way that the control obtained takes into account the dual objective of minimizing the objective functional J as well as minimizing the sensitivity of J with respect to changes in the coefficients. For illustration, two numerical examples are included.

2. Problem statement

This article is concerned with the sensitivity of the following unconstrained optimal control problem, which is denoted Problem (P):

Subject to the dynamical system

$$\dot{x}(t) = f(t, x(t), u(t), a), \tag{1a}$$

$$x(0) = x_0(a), \tag{1b}$$

*This paper was partially supported by a research grant from the Australian Research Council.

find a control $u \in \mathcal{U}$ such that

$$J = \Phi_0(x(T), a) + \int_0^T \mathcal{L}_0(t, x(t), u(t), a)dt \tag{2}$$

is minimized over $\mathcal{U}$, where $x : [0, T] \to \mathbb{R}^n$ is called the state; $u : [0, T] \to \mathbb{R}^r$ is called the control; $a \in \mathbb{R}^m$ is called the coefficient vector; $f: [0, T] \times \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^m \to \mathbb{R}^n$; $\mathcal{L}_0 : [0, T] \times \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^m \to \mathbb{R}$; $\Phi_0 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$; $x_0 : \mathbb{R}^m \to \mathbb{R}^n$; and

$$\mathcal{U} = \{u \in L_\infty^r : \underline{u}_i \leq u_i(t) \leq \bar{u}_i \text{ a.e. in } [0, T], \underline{u}_i, \bar{u}_i \in \mathbb{R}; i \in [1, \ldots, r]\}$$

is known as the set of admissible controls, where L_∞^r denotes the Banach space $L_\infty([0, T], \mathbb{R}^r)$ of all essentially bounded measurable functions from $[0, T]$ into $\mathbb{R}^r$. Its norm is

$$\|u\|_\infty = \operatorname*{ess\,sup}_{t \in [0,T]} \left\{ \sum_{i=1}^r (u_i(t))^2 \right\}^{1/2}.$$

If the coefficient vector a is specified, then Problem (P) may be solved numerically by several methods such as the gradient-restoration algorithms [5] and the control parameterization technique [1, 7]. Note, that the control parameterization technique involves partitioning the interval $[0, T]$ and then approximating the controls by piecewise constant functions that are consistent with the partitioning. The reduced problem can then be viewed as a mathematical programming problem.

In more detail, for $i = 1, 2, \ldots, r$, let $\{\mathcal{I}_i^p\}_{p=1}^\infty$ be a sequence of partitions of the interval $[0, T]$ such that $\mathcal{I}_i^p$ has n_p elements, $\mathcal{I}_i^{p+1}$ is a refinement of $\mathcal{I}_i^p$, and $\|\mathcal{I}_i^p\| \to 0$ as $p \to \infty$, where $\|\mathcal{I}_i^p\|$ denotes the length of the largest interval in the partition $\mathcal{I}_i^p$. We assume that

$$\mathcal{I}_i^p = \{\mathcal{I}_{ij}^p\}_{j=1}^{n_p},$$

where $I_{ij}^p = [t_{i,j-1}^p, t_{ij}^p]$ and $0 = t_{i0}^p < t_{i1}^p < \ldots < t_{in_p}^p = T$. Then,

$$\|\mathcal{I}_i^p\| = \max_{1 \leq j \leq n_p} l(I_{ij}^p),$$

where $l(I_{ij}^p) = t_{ij}^p - t_{i,j-1}^p$.

Let $\mathcal{U}^p$ be the subset of admissible controls which are piecewise constant and consistent with the partitions $\mathcal{I}_i^p$, $i = 1, 2, \ldots, r$. Clearly, each $u^p \in \mathcal{U}^p$ can be written as

$$u_i^p(t) = \sum_{k=1}^{n_p} \sigma_{ik}^p \chi_{I_{ik}^p}(t), \qquad t \in [0, T], \ i = 1, 2, \ldots, r \tag{3}$$

where $\sigma_{ik}^p \in \mathbb{R}$ and $\chi_{I_{ik}^p}(t)$ denotes the indicator function of I_{ik}^p, defined by

$$\chi_I(t) = \begin{cases} 1, & \text{if } t \in I \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Then, each control $u^p \in \mathcal{U}^p$ can be identified uniquely with a control parameter vector σ^p and vice versa, where

$$\sigma^p = [(\sigma_1^p)^{\mathrm{T}}, (\sigma_2^p)^{\mathrm{T}}, \ldots, (\sigma_r^p)^{\mathrm{T}}]^{\mathrm{T}},$$

and

$$\sigma_i^p = [\sigma_{i1}^p, \sigma_{i2}^p, \sigma_{i3}^p, \ldots, \sigma_{in_p}^p]^{\mathrm{T}} \qquad , i = 1, 2, \ldots, r.$$

By virtue of the definition of $\mathcal{U}$, the control parameters are subject to the following constraints:

$$\underline{u}_i \leq \sigma_{ij}^p \leq \bar{u}_i, \qquad i = 1, 2, \ldots, r; j = 1, 2, \ldots, n_p. \tag{5}$$

Denote by Ξ^p the subset of $\mathbb{R}^{n_p r}$ which consists of all vectors σ^p such that the constraints (5) are satisfied. Ξ^p is called the set of admissible control parameters.

Using this approach, we obtain the following approximate problem, which is referred to as Problem Q(p).

Subject to the dynamical system

$$\dot{x}(t) = \hat{f}(t, x(t), \sigma^p, a), \tag{6a}$$

$$x(0) = x_0(a), \tag{6b}$$

find a control parameter $\sigma^p \in \Xi^p$ such that

$$\hat{J}^p = \Phi_0(x(T|\sigma^p), a) + \int_0^T \hat{\mathcal{L}}_0(t, x(t|\sigma^p), \sigma^p, a)\, dt \tag{7}$$

is minimized over Ξ^p, where $x(\cdot|\sigma^p)$ denotes the solution of the system (1) when u^p is given by equation (3), $\hat{\mathcal{L}}_0(t, x(t), \sigma^p, a) = \mathcal{L}_0\big(t, x(t), \Sigma_{k=1}^{n_p} \sigma_{ik}^p \chi_{ik}^p(t), a\big)$, and $\hat{f}(t, x(t), \sigma^p, a) = f\big(t, x(t), \Sigma_{k=1}^{n_p} \sigma_{ik}^p \chi_{ik}^p(t), a\big)$.

Define the Hamiltonian for Problem (P) as:

$$H(t, x(t), u(t), \lambda(t), a) = \mathcal{L}_0(t, x(t), u(t), a) + [\lambda(t)]^{\mathrm{T}} f(t, x(t), u(t), a), \tag{8}$$

where λ is the costate vector satisfying:

$$[\dot{\lambda}(t)]^{\mathrm{T}} = -\frac{\partial H(t, x(t), u(t), \lambda(t), a)}{\partial x}, \tag{9a}$$

$$[\lambda(T)]^{\mathrm{T}} = \frac{\partial \Phi_0(x(T), a)}{\partial x(T)}. \tag{9b}$$

The gradient formula for the objective functional (7) is well known (c.f.[2]). It is given in the following theorem.

Theorem 1. Consider Problem (Q(p)). The gradient of $\hat{J}^p$ with respect to control parameter σ_{ij}^p is given by

$$\frac{\partial \hat{J}^p}{\partial \sigma_{ij}^p} = \int_{t_{ij-1}^p}^{t_{ij}^p} \frac{\partial H(t, x(t|\sigma^p), u^p(t), \lambda(t|\sigma^p), a)}{\partial u_i} dt, \tag{10}$$

where u^p is defined by equation (3), and $\lambda(\cdot|\sigma^p)$ denotes the solution of the costate system corresponding to u^p.

Using the result of Theorem 1, the gradient of (7) corresponding to each control parameter $\sigma^p \in \Xi^p$ may be calculated using the following plan, which is referred to as Plan 1.

1. Integrate the state equations (1a) together with the initial conditions (1b) forward in time from $t = 0$ to $t = T$.
2. Integrate the costate equations (9a) together with the final conditions (9b) backward in time from $t = T$ to $t = 0$.
3. Calculate the required gradient using the formula (10)

With the aid of Plan 1, Problem (Q(p)) may be solved by using standard mathematical programming techniques.

We may now formulate a new problem so as to take into account the sensitivity of the optimum cost with respect to changes in the coefficient vector a. It is done by incorporating the function

$$\left(\frac{\partial J}{\partial a}\right) \left(\frac{\partial J}{\partial a}\right)^{\mathrm{T}}.$$

in the objective functional. In this process, we obtain a new problem, which is referred to as Problem (S).

Subject to the dynamical system

$$\dot{x}(t) = f(t, x(t), u(t), a), \tag{11a}$$

$$x(0) = x_0(a), \tag{11b}$$

find a control $u \in \mathcal{U}$ such that

$$G = J + \alpha \left(\frac{\partial J}{\partial a}\right) \left(\frac{\partial J}{\partial a}\right)^{\mathrm{T}} \tag{12}$$

is minimized over $\mathcal{U}$, where x, u, a, Φ_0, $\mathcal{L}_0$, f, x_0 and $\mathcal{U}$ are all as defined in Problem (P), α is a weighting factor, and J is the original objective functional

$$J = \Phi_0(x(T), a) + \int_0^T \mathcal{L}_0(t, x(t), u(t), a)\, dt. \tag{13}$$

Problem (S) is of a different nature than Problem (P). Plan 1 cannot be used to find the gradient of the objective functional (12) in the parameterized version of Problem (S) because of the term $\left(\frac{\partial J}{\partial a}\right)\left(\frac{\partial J}{\partial a}\right)^{\mathrm{T}}$ appearing in the new objective functional. Problem (S) needs to be handled differently and the details are given in the next section.

3. Solution method

Let the control parameters σ^p and the set Ξ^p be as defined in Section 2. Now, by applying the technique of control parameterization to Problem (S), we obtain the approximate problem, Problem (T(p)):

Subject to the dynamical system

$$\dot{x}(t) = \hat{f}(t, x(t), \sigma^p, a), \tag{14a}$$

$$x(0) = x_0(a) \tag{14b}$$

find a control parameter $\sigma^p \in \Xi^p$ such that

$$\hat{G}^p = \hat{J}^p + \alpha \left(\frac{\partial \hat{J}^p}{\partial a}\right) \left(\frac{\partial \hat{J}^p}{\partial a}\right)^{\mathrm{T}} \tag{15}$$

is minimized over Ξ^p, where

$$\hat{J}^p = \Phi_0(x(T|\sigma^p), a) + \int_0^T \hat{\mathfrak{L}}_0(t, x(t|\sigma^p), \sigma^p, a)\, dt, \tag{16}$$

and $\hat{\mathfrak{L}}_0$, $\hat{f}$, and $x(\cdot|\sigma^p)$ have the same meaning as in Problem (Q(p)).

To solve Problem (T(p)), we need the gradient formula for the objective functional (15). Let us establish a number of preliminary results. Since the notation in this section will soon become very cumbersome, we will simplify it somewhat by temporarily abandoning the 'hat' notation in the definition of Problem (T(p)) and by writing σ for σ^p.

Lemma 1. The gradient of the objective functional (16) with respect to a is:

$$\frac{\partial J}{\partial a} = \frac{\partial \Phi_0(x(T|\sigma), a)}{\partial a} + [\lambda(0)|\sigma)]^{\mathrm{T}} \frac{\partial x_0(a)}{\partial a} + \int_0^T \frac{\partial H}{\partial a}\, dt, \tag{17}$$

where

$$H(t, x(t), \sigma, \lambda(t), a) = \mathfrak{L}_0(t, x(t), \sigma, a) + [\lambda(t)]^{\mathrm{T}} f(t, x(t), \sigma, a), \tag{18}$$

and $\lambda(\cdot|\sigma)$ is the solution of the costate system:

$$[\dot{\lambda}(t)]^{\mathrm{T}} = -\frac{\partial H(t, x(t), \sigma, \lambda(t), a)}{\partial x} \tag{19a}$$

$$[\lambda(T)]^{\mathrm{T}} = \frac{\partial \Phi_0(x(T), a)}{\partial x(T)}. \tag{19b}$$

corresponding to $\sigma \in \Xi^p$.

Proof. Let $a_0 \in \mathbb{R}^m$ be given and $\rho_0 \in \mathbb{R}^m$ be arbitrary but fixed. Define

$$a(\epsilon) = a_0 + \epsilon\rho_0, \tag{20}$$

where $\epsilon > 0$ is arbitrarily small. Let $x(\cdot)$ and $x(\cdot; \epsilon)$ denote, respectively, the solution of the dynamical system (14) corresponding to a_0 and $a(\epsilon)$. Then, from (14),

$$x(t) = x_0(a) + \int_0^t f(s, x(s), \sigma, a_0)ds \tag{21}$$

$$x(t; \epsilon) = x_0(a(\epsilon)) + \int_0^t f(s, x(s; \epsilon), \sigma, a(\epsilon))ds. \tag{22}$$

Hence, by Taylor's Theorem, it follows that

$$\begin{aligned}\delta x(t) &\equiv \lim_{\epsilon \to 0}\left\{\frac{x(t; \epsilon) - x(t)}{\epsilon}\right\} \\ &= \frac{\partial x_0(a)}{\partial a}\rho_0 + \int_0^t \left\{\frac{\partial f(s, x(s), \sigma, a_0)}{\partial x}\delta x(s) + \frac{\partial f(s, x(s), \sigma, a_0)}{\partial a}\rho_0\right\}ds. \end{aligned} \tag{23}$$

Clearly,

$$\delta\dot{x}(t) = \frac{\partial f(t, x(t), \sigma, a_0)}{\partial x}\delta x(t) + \frac{\partial f(t, x(t), \sigma, a_0)}{\partial a}\rho_0 \tag{24}$$

$$\delta x(0) = \frac{\partial x_0(a)}{\partial a}\rho_0. \tag{25}$$

Now, we express J as

$$J(a) = \Phi_0(x(T), a) + \int_0^T \{H(t, x(t), \sigma, \lambda(t), a) - [\lambda(t)]^{\mathrm{T}} f(t, x(t), \sigma, a)\}dt. \tag{26}$$

Then it follows that

$$\begin{aligned}\delta J(a) &= \frac{\partial J(a_0)}{\partial a}\rho_0\\&= \frac{\partial \Phi_0(x(T), a)}{\partial x}\delta x(T) + \frac{\partial \Phi_0(x(T), a)}{\partial a}\rho_0\\&\quad + \int_0^T \{\delta H(t, x(t), \sigma, \lambda(t), a_0) - [\lambda(t)]^{\mathrm{T}}\delta\dot{x}(t) - [\delta\lambda(t)]^{\mathrm{T}} f(t, x(t), \sigma, a_0)\}dt,\end{aligned} \tag{27}$$

where

$$\begin{aligned}\delta H(t, x(t), \sigma, \lambda(t), a_0) &= \frac{\partial H(t, x(t), \sigma, \lambda(t), a_0)}{\partial x}\delta x(t)\\&\quad + \frac{\partial H(t, x(t), \sigma, \lambda(t), a_0)}{\partial a}\rho_0\\&\quad + \frac{\partial H(t, x(t), \sigma, \lambda(t), a_0)}{\partial \lambda}\delta\lambda(t).\end{aligned} \tag{28}$$

From the definition of H, we have

$$\frac{\partial H(t, x(t), \sigma, \lambda(t), a_0)}{\partial \lambda} = [f(t, x(t), \sigma, a_0)]^{\mathrm{T}} = [\dot{x}(t)]^{\mathrm{T}}. \tag{29}$$

Substituting equations (19a) and (29) into equation (28), we get

$$\begin{aligned}\delta H(t, x(t), \sigma, \lambda(t), a_0) &= -[\dot{\lambda}(t)]^{\mathrm{T}}\delta x(t) + [\dot{x}(t)]^{\mathrm{T}}\delta\lambda(t)\\&\quad + \frac{\partial H(t, x(t), \sigma, \lambda(t), a_0)}{\partial a}\rho_0.\end{aligned} \tag{30}$$

Then, equation (27) becomes

$$\begin{aligned}\frac{\partial J(a_0)}{\partial a}\rho_o &= \frac{\partial \Phi_0(x(T), a_0)}{\partial x}\delta x(T) + \frac{\partial \Phi_0(x(T), a_0)}{\partial a}\rho_o\\&\quad + \int_0^T \left\{-\frac{d[[\lambda(t)]^{\mathrm{T}}\delta x(t)]}{dt} + \frac{\partial H(t, x(t), \sigma, \lambda(t), a_0)}{\partial a}\rho_0\right\}dt\\&= \frac{\partial \Phi_0(x(T), a_0)}{\partial x}\delta x(T) + \frac{\partial \Phi_0(x(T), a_0)}{\partial a}\rho_0 - [\lambda(T)]^{\mathrm{T}}\delta x(T)\\&\quad + [\lambda(0)]^{\mathrm{T}}\delta x(0) + \int_0^T \left\{\frac{\partial H(t, x(t), \sigma, \lambda(t), a_0)}{\partial a}\rho_0\right\}dt.\end{aligned} \tag{31}$$

Using equations (19b) and (25), (31) becomes

$$\frac{\partial J(a_0)}{\partial a}\rho_0 = \frac{\partial \Phi_0(x(T), a_0)}{\partial a}\rho_0 + [\lambda(0)]^{\mathrm{T}} \frac{\partial x_0(a)}{\partial a}\rho_0 + \int_0^T \left\{ \frac{\partial H(t, x(t), \sigma, \lambda(t), a_0)}{\partial a}\rho_0 \right\} dt. \tag{32}$$

Since ρ_0 is arbitrary, it follows that

$$\frac{\partial J}{\partial a} = \frac{\partial \Phi_0(x(T), a)}{\partial a} + [\lambda(0)]^{\mathrm{T}} \frac{\partial x_0(a)}{\partial a} + \int_0^T \frac{\partial H}{\partial a} dt,$$

as required.

Lemma 2. The gradient of the objective functional (16) with respect to σ_{ij} is given by

$$\frac{\partial J}{\partial \sigma_{ij}} = \int_0^T \frac{\partial H}{\partial \sigma_{ij}} dt. \tag{33}$$

Proof. Since σ_{ij} can be treated similarly to a, the proof is almost identical to that of Lemma 1, except that Φ_0 as well as x_0 are not an explicit function of σ_{ij}. Also σ_{ij} is not a vector.

Lemma 3. Let $\phi : [0, T] \to \mathbb{R}^n \times \mathbb{R}^m$; $\psi : [0, T] \to \mathbb{R}^n \times \mathbb{R}^m$; and $\bar{H}^{\mathrm{T}} \in \mathbb{R}^m$ be such that

$$\bar{H} = \frac{\partial H}{\partial a} + f^{\mathrm{T}} \phi + \left(\frac{\partial H}{\partial x} \right) \psi, \tag{34}$$

$$\dot{\phi}(t) = - \left(\frac{\partial (\bar{H}^{\mathrm{T}})}{\partial x} \right)^{\mathrm{T}}, \tag{35a}$$

$$\phi(T) = \left(\frac{\partial}{\partial x(T)} \left(\frac{\partial \Phi_0}{\partial a} \right)^{\mathrm{T}} \right)^{\mathrm{T}}, \tag{35b}$$

$$\dot{\psi}(t) = \left(\frac{\partial (\bar{H}^{\mathrm{T}})}{\partial \lambda} \right)^{\mathrm{T}}, \tag{36a}$$

$$\psi(0) = \frac{\partial x_0(a)}{\partial a}. \tag{36b}$$

Then,

$$\frac{\partial}{\partial \sigma_{ij}} \left(\frac{\partial J}{\partial a} \right) = \int_0^T \frac{\partial \bar{H}}{\partial \sigma_{ij}} dt. \tag{37}$$

Proof. Note that both x and λ depend on σ. However, for convenience, we shall use $x(\cdot)$ and $\lambda(\cdot)$ to denote $x(\cdot|\sigma)$ and $\lambda(\cdot|\sigma)$, respectively. From Lemma 1, we have

$$\frac{\partial}{\partial\sigma_{ij}}\left(\frac{\partial J}{\partial a}\right) = \frac{\partial}{\partial\sigma_{ij}}\left(\int_0^T\frac{\partial H}{\partial a}dt + \frac{\partial\Phi_0(x(T),a)}{\partial a} + [\lambda(0)]^{\mathrm{T}}\frac{\partial x_0(a)}{\partial a}\right).$$

Let

$$\bar{G}_a = \int_0^T\left\{\frac{\partial H}{\partial a} + (f-\dot{x})^{\mathrm{T}}\phi + \left[\frac{\partial H}{\partial x} + \dot{\lambda}^{\mathrm{T}}\right]\psi\right\}dt + \frac{\partial\Phi_0(x(T),a)}{\partial a} + [\lambda(0)]^{\mathrm{T}}\frac{\partial x_0(a)}{\partial a}. \tag{38}$$

Then,

$$\bar{G}_a = \int_0^T\{\bar{H} - \dot{x}^{\mathrm{T}}\phi + \dot{\lambda}^{\mathrm{T}}\psi\}\,dt + \frac{\partial\Phi_0(x(T),a)}{\partial a} + [\lambda(0)]^{\mathrm{T}}\frac{\partial x_0(a)}{\partial a}. \tag{39}$$

Integrating by parts, equation (39) becomes

$$\begin{aligned}\bar{G}_a &= \int_0^T\bar{H}\,dt - [x^{\mathrm{T}}\phi]_0^T + \int_0^T x^{\mathrm{T}}\dot{\phi}\,dt + [\lambda^{\mathrm{T}}\psi]_0^T - \int_0^T\lambda^{\mathrm{T}}\dot{\psi}\,dt \\ &\quad + \frac{\partial\Phi_0(x(T),a)}{\partial a} + [\lambda(0)]^{\mathrm{T}}\frac{\partial x_0(a)}{\partial a} \\ &= \int_0^T\{\bar{H} + x^{\mathrm{T}}\dot{\phi} - \lambda^{\mathrm{T}}\dot{\psi}\}\,dt - x^{\mathrm{T}}(T)\phi(T) + x^{\mathrm{T}}(0)\phi(0) \\ &\quad + \lambda^{\mathrm{T}}(T)\psi(T) - \lambda^{\mathrm{T}}(0)\psi(0) + \frac{\partial\Phi_0(x(T),a)}{\partial a} + [\lambda(0)]^{\mathrm{T}}\frac{\partial x_0(a)}{\partial a}\end{aligned} \tag{40}$$

Now, a small perturbation in σ_{ij} will cause a perturbation in x and λ and both of these will induce a perturbation in $\bar{G}_a$. Hence, by using the chain rule and equation (40), we get

$$\begin{aligned}\delta\bar{G}_a &= \int_0^T[\delta x(t)]^{\mathrm{T}}\left[\left(\frac{\partial(\bar{H}^{\mathrm{T}})}{\partial x}\right)^{\mathrm{T}} + \dot{\phi}\right]dt + \int_0^T[\delta\lambda(t)]^{\mathrm{T}}\left[\left(\frac{\partial(\bar{H}^{\mathrm{T}})}{\partial\lambda}\right)^{\mathrm{T}} - \dot{\psi}\right]dt \\ &\quad + \int_0^T\frac{\partial\bar{H}}{\partial\sigma_{ij}}\delta\sigma_{ij}dt - [\delta x(T)]^{\mathrm{T}}\phi(T) + [\delta x(0)]^{\mathrm{T}}\phi(0) + [\delta\lambda(T)]^{\mathrm{T}}\psi(T) \\ &\quad - [\delta\lambda(T)]^{\mathrm{T}}\psi(T) + [\delta x(T)]^{\mathrm{T}}\left[\frac{\partial}{\partial x(T)}\left(\frac{\partial\Phi_0}{\partial a}\right)^{\mathrm{T}}\right]^{\mathrm{T}} + [\delta\lambda(0)]^{\mathrm{T}}\frac{\partial x_0(a)}{\partial a}.\end{aligned} \tag{41}$$

Both the initial conditions for the states and the final conditions for the costates are prescribed and fixed so that

$$\delta x(0) = 0$$
$$\delta\lambda(T) = 0.$$

Hence, equation (41) becomes

$$\delta\bar{G}_a = \int_0^T [\delta x(t)]^{\mathrm{T}} \left[\left(\frac{\partial(\bar{H}^{\mathrm{T}})}{\partial x} \right)^{\mathrm{T}} + \dot{\phi} \right] dt + \int_0^T [\delta\lambda(t)]^{\mathrm{T}} \left[\left(\frac{\partial(\bar{H}^{\mathrm{T}})}{\partial \lambda} \right)^{\mathrm{T}} - \dot{\psi} \right] dt$$
$$+ \delta\sigma_{ij} \int_0^T \frac{\partial \bar{H}}{\partial \sigma_{ij}} dt + [\delta x(T)]^{\mathrm{T}} \left[\left(\frac{\partial}{\partial x(T)} \left(\frac{\partial \Phi_0}{\partial a} \right)^{\mathrm{T}} \right)^{\mathrm{T}} - \phi(T) \right]$$
$$+ [\delta\lambda(0)]^{\mathrm{T}} \left[\frac{\partial x_0(a)}{\partial a} - \psi(0) \right]. \tag{42}$$

From equations (35) and (36), equation (42) simplifies to

$$\delta\bar{G}_a = \delta\sigma_{ij} \int_0^T \frac{\partial \bar{H}}{\partial \sigma_{ij}} dt. \tag{43}$$

Thus, it follows that

$$\frac{\partial}{\partial \sigma_{ij}} \left(\frac{\partial J}{\partial a} \right) = \int_0^T \frac{\partial \bar{H}}{\partial \sigma_{ij}} dt.$$

This completes the proof.

Remark 1. Returning to the 'hat' notation, we have

$$\hat{H}(t, x(t), \sigma, \lambda(t), a) = H(t, x(t), u(t), \lambda(t), a).$$

So,

$$\frac{\partial \hat{H}}{\partial \sigma_{ij}} = \frac{\partial H}{\partial u_i} \frac{\partial u_i}{\partial \sigma_{ij}}$$

where, from equation (3),

$$\frac{\partial u_i}{\partial \sigma_{ij}} = \begin{cases} 1, & t_{ij-1} \leq t \leq t_{ij} \\ 0, & \text{otherwise.} \end{cases}$$

Hence, Lemma 2 becomes

$$\frac{\partial \hat{J}}{\partial \sigma_{ij}} = \int_{t_{ij-1}}^{t_{ij}} \frac{\partial H}{\partial u_i} dt. \tag{44}$$

By a similar argument for $\hat{\bar{H}}$ and $\bar{H}$, Lemma 3 becomes

$$\frac{\partial}{\partial \sigma_{ij}} \left(\frac{\partial \hat{J}}{\partial a} \right) = \int_{t_{ij-1}}^{t_{ij}} \frac{\partial \bar{H}}{\partial u_i} dt \tag{45}$$

We are now in a position to present the gradient formula for the objective functional (15) in the following theorem.

Theorem 2. The gradient of the objective functional (15) is:

$$\frac{\partial \hat{G}}{\partial \sigma_{ij}} = \int_{t_{ij-1}}^{t_{ij}} \frac{\partial H}{\partial u_i} dt + 2\alpha \left[\int_{t_{ij-1}}^{t_{ij}} \frac{\partial \bar{H}}{\partial u_i} dt \right] \left[\frac{\partial \Phi_0(x(T|\sigma), a)}{\partial a} + [\lambda(0)|\sigma)]^{\mathrm{T}} \frac{\partial x_0(a)}{\partial a} + \int_0^T \frac{\partial H}{\partial a} dt \right]^{\mathrm{T}} \tag{46}$$

where $\hat{G}$ is given by equation (15); H is given by equation (18); and $\bar{H}$ is given by equation (34).

Proof. We omit the 'hat' notation here for clarity. First,

$$\frac{\partial G}{\partial \sigma_{ij}} = \frac{\partial}{\partial \sigma_{ij}} \left(J + \alpha \left(\frac{\partial J}{\partial a} \right) \left(\frac{\partial J}{\partial a} \right)^{\mathrm{T}} \right)$$

$$= \frac{\partial J}{\partial \sigma_{ij}} + 2\alpha \left(\frac{\partial}{\partial \sigma_{ij}} \left(\frac{\partial J}{\partial a} \right) \right) \left(\frac{\partial J}{\partial a} \right)^{\mathrm{T}}. \tag{47}$$

Using Lemmas 1, 2, and 3 and Remark 1, equation (47) becomes

$$\frac{\partial G}{\partial \sigma_{ij}} = \int_{t_{ij-1}}^{t_{ij}} \frac{\partial H}{\partial u_i} dt + 2\alpha \left(\int_{t_{ij-1}}^{t_{ij}} \frac{\bar{H}}{\partial u_i} dt \right) \left(\frac{\partial J}{\partial a} \right)^{\mathrm{T}}$$

$$= \int_{t_{ij-1}}^{t_{ij}} \frac{\partial H}{\partial u_i} dt + 2\alpha \left(\int_{t_{ij-1}}^{t_{ij}} \frac{\partial \bar{H}}{\partial u_i} dt \right) \left(\frac{\partial \Phi_0(x(T|\sigma), a)}{\partial a} + [\lambda(0|\sigma)]^{\mathrm{T}} \frac{\partial x_0(a)}{\partial a} + \int_0^T \frac{\partial H}{\partial a} dt \right)^{\mathrm{T}}.$$

Consequently, the result follows readily.

Looking at equation (36), we see that it is reducible to:

$$\dot{\psi}(t) = \frac{\partial f}{\partial a} + \frac{\partial f}{\partial a}\psi(t) \tag{48a}$$

$$\psi(0) = \frac{\partial x_0(a)}{\partial a}. \tag{48b}$$

The system (48) is independent of both λ and ϕ. Thus, it is solvable without knowing either of these. Next, let us examine the right hand side of equation (35a) and the integrand of equation (45). We have

$$\begin{aligned}\frac{\partial \bar{H}}{\partial u_i} &= \frac{\partial}{\partial u_i}\left(\frac{\partial H}{\partial a}\right) + \left(\frac{\partial f}{\partial u_i}\right)^{\mathsf{T}}\phi + \left(\frac{\partial}{\partial u_i}\left(\frac{\partial H}{\partial x}\right)\right)\psi \\ &= \frac{\partial}{\partial u_i}\left(\frac{\partial \mathfrak{L}_0}{\partial a} + \lambda^{\mathsf{T}}\frac{\partial f}{\partial a}\right) + \left(\frac{\partial f}{\partial u_i}\right)^{\mathsf{T}}\phi + \left(\frac{\partial}{\partial u_i}\left(\frac{\partial \mathfrak{L}_0}{\partial x} + \lambda^{\mathsf{T}}\frac{\partial f}{\partial x}\right)\right)\psi \\ &= \frac{\partial^2 \mathfrak{L}_0}{\partial u_i \partial a} + \lambda^{\mathsf{T}}\frac{\partial^2 f}{\partial u_i \partial a} + \left(\frac{\partial f}{\partial u_i}\right)^{\mathsf{T}}\phi + \left(\frac{\partial^2 \mathfrak{L}_0}{\partial u_i \partial x} + \lambda^{\mathsf{T}}\frac{\partial^2 f}{\partial u_i \partial x}\right)\psi, \text{ for } i = 1, \ldots, r.\end{aligned} \tag{49}$$

Similarly,

$$\frac{\partial \bar{H}}{\partial x_i} = \frac{\partial^2 \mathfrak{L}_0}{\partial x_i \partial a} + \lambda^{\mathsf{T}}\frac{\partial^2 f}{\partial x_i \partial a} + \left(\frac{\partial f}{\partial x_i}\right)^{\mathsf{T}}\phi + \left(\frac{\partial^2 \mathfrak{L}_0}{\partial x_i \partial x} + \lambda^{\mathsf{T}}\frac{\partial^2 f}{\partial x_i \partial x}\right)\psi, \text{ for all } i = 1, \ldots, n. \tag{50}$$

The system (35) becomes

$$\dot{\phi}(t) = -\left(\frac{\partial(\bar{H}^{\mathsf{T}})}{\partial x}\right)^{\mathsf{T}} = -\begin{bmatrix} \frac{\partial \bar{H}_1}{\partial x_1} & \frac{\partial \bar{H}_2}{\partial x_1} & \cdots & \frac{\partial \bar{H}_m}{\partial x_1} \\ \frac{\partial \bar{H}_1}{\partial x_2} & \frac{\partial \bar{H}_2}{\partial x_2} & \cdots & \frac{\partial \bar{H}_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \bar{H}_1}{\partial x_n} & \frac{\partial \bar{H}_2}{\partial x_n} & \cdots & \frac{\partial \bar{H}_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} -\frac{\partial(\bar{H}^{\mathsf{T}})}{\partial x_1} \\ -\frac{\partial(\bar{H}^{\mathsf{T}})}{\partial x_2} \\ \vdots \\ -\frac{\partial(\bar{H}^{\mathsf{T}})}{\partial x_n} \end{bmatrix}, \tag{51a}$$

$$\phi(T) = \left(\frac{\partial}{\partial x(T)}\left(\frac{\partial \Phi_0}{\partial a}\right)^{\mathsf{T}}\right)^{\mathsf{T}}, \tag{51b}$$

where $\frac{\partial \bar{H}}{\partial x_i}$, $i = 1, \ldots, n$, are given by equation (50).

We are now in a position to give a plan, denoted by Plan 2, for finding the gradient of the objective functional (15):

1. Integrate the state equations (11a) together with the initial conditions (11b) forward in time from $t = 0$ to $t = T$.
2. Integrate the costate equations (19a) together with the final conditions (19b) backward in time from $t = T$ to $t = 0$.
3. Calculate $\frac{\partial \hat{J}}{\partial \sigma_{ij}}$ from (44), using $\frac{\partial H}{\partial u_i}$ for $i = 1, \ldots, r$.
4. Calculate $\frac{\partial \hat{J}}{\partial a}$ from (17), using $\frac{\partial H}{\partial a}$.
5. Integrate the ψ equations (48a) together with the initial conditions (48b) forward in time from $t = 0$ to $t = T$.
6. Integrate the ϕ equations (51a) together with the final conditions (51b) backward in time from $t = T$ to $t = 0$.
7. Calculate $\frac{\partial \hat{G}}{\partial \sigma_{ij}}$ from (46), using $\frac{\partial \bar{H}}{\partial u_i}$ for $i = 1, \ldots, r$ as given by (49).

Using this plan, we may, as before, apply the techniques of mathematical programming to solve Problem (T(p)). From now on we will no longer use the 'hat' notation, so that $\hat{J}$ and $\hat{G}$ are referred to as J and G respectively.

4. Examples

The general purpose software MISER3 [3] is an efficient one for solving Problem (P) via the control parameterization technique. However, it cannot be used directly for solving optimal control problems involving the cost functional of the form (15). On the basis of Plan 2, we have modified the code of MISER3 for our purpose. The new code is used to solve the following two examples.

Example 1.[6]. Consider the following dynamical system:

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= -x_1(t) + (1.4 - 0.14(x_2(t))^2)x_2(t) + 4u(t), \end{aligned} \tag{52a}$$

with initial conditions

$$\begin{aligned} x_1(0) &= -5 + a \\ x_2(0) &= -5. \end{aligned} \tag{52b}$$

The objective functional is:

$$J = \frac{1}{2}\int_0^{10} ((x_1(t))^2 + (u(t))^2)dt. \tag{53}$$

Here, we have one coefficient a which occurs only in the initial conditions and we set $a = 0$. We solve two problems. The first is just the standard optimal control problem (52)–(53). The second problem is of the form (S), which is obtained from the problem (52)–(53) by incorporating the sensitivity factor in the objective functional. More precisely, the new problem is: subject to the dynamical system (52), find a control u, such that the "objective functional"

$$G = J + \alpha \left(\frac{\partial J}{\partial a}\right)^2 \tag{54}$$

is minimized, where J is defined by (53) and α is the weighting factor. In both cases, n_p is taken as 20. The weighting factor α is set to be 1. Let u^* and $\bar{u}$ be, respectively, the optimal controls of the problems (52) with equation (53) and (52) with equation (54), both with $a = 0$. These are shown in Figure 1. Furthermore, let $J(u^*)$ and $J(\bar{u})$ denote, respectively, the values of J corresponding to the controls u^* and $\bar{u}$. As we expect, $J(u^*) = 15.882 < J(\bar{u}) = 17.948$. However, the value of $\|\frac{\partial J}{\partial a}\|$, which gives an idea of the sensitivity of J with respect to a, has been reduced significantly when $\bar{u}$ instead of u^* is used. This is shown in Table 1.

Table 1. Optimal solution for Example 1.

Problem	Objective J	$\left\|\frac{\partial J}{\partial a}\right\|$
(52) with (53)	15.882	4.5464
(52) with (54)	17.948	0.0312

Let us now see how varying the value of a about $a = 0$ influences the value of J in both cases. The controls obtained from $a = 0$, u^* and $\bar{u}$, are taken to be the base controls. J is recalculated using the base controls and changing the values of a for both the original and the new problem. The results are shown in Figure 2. The first thing to note is that $J(u^*)$ is very sensitive with respect to changes in a. For example, a change of 0.1 in the value of a causes $J(u^*)$ to nearly double. The change in $J(\bar{u})$, however, is much smaller. For most values of a, except those near $a = 0$, $J(\bar{u})$ is far superior to $J(u^*)$. A few numerical values to compare the two solutions are given in Table 2.

Example 2.[4]. Consider the following dynamical system:

$$\dot{x}_1(t) = -(x_1(t) + 0.25)(2.0 + u(t)) + (x_2(t) + 0.5)e(x_1(t))$$

$$\dot{x}_2(t) = 0.5 - x_2(t) - (x_2(t) + 0.5)e(x_1(t)), \tag{55a}$$

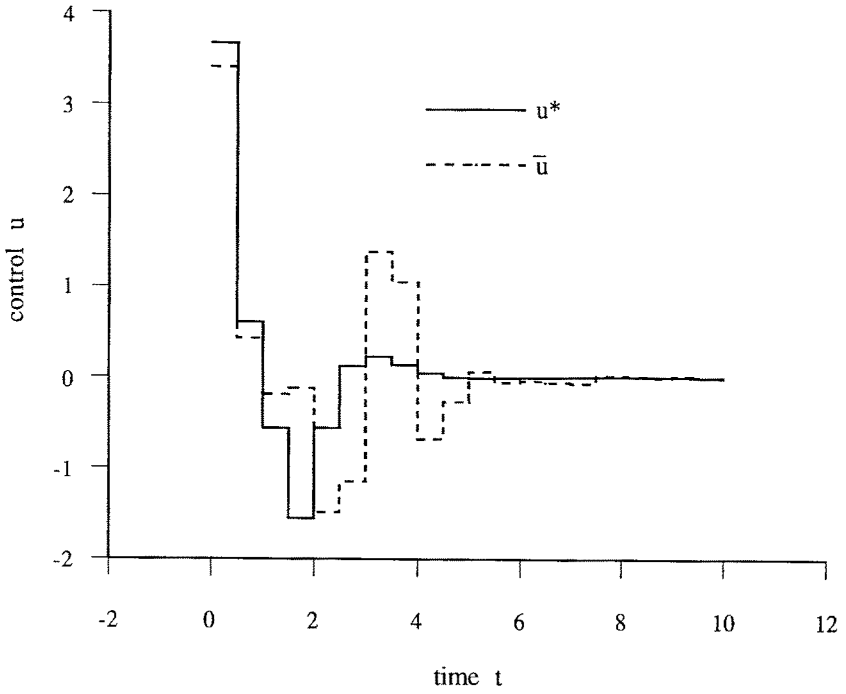

Figure 1. Optimal controls computed for Example 1.

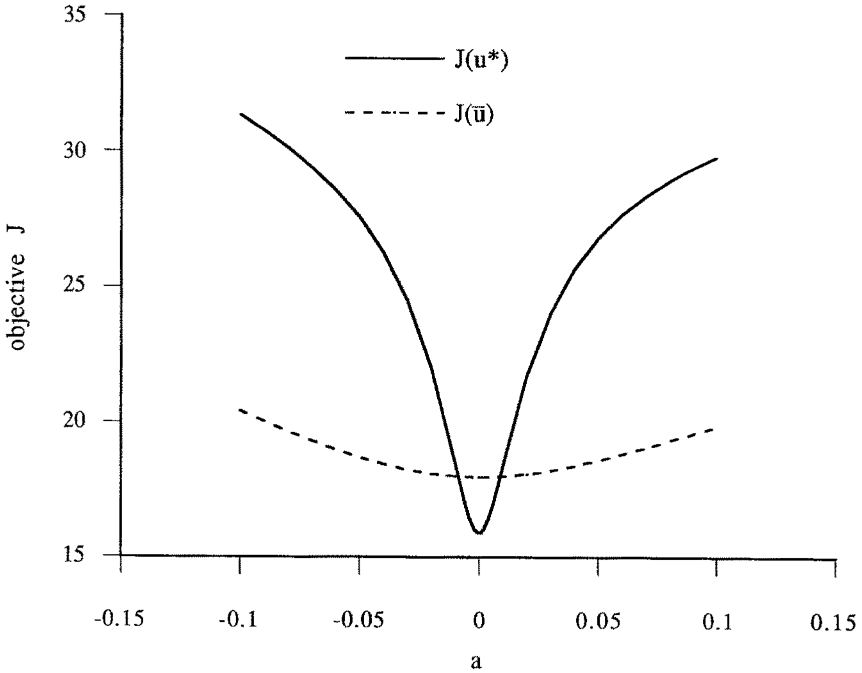

Figure 2. Variation in objective functional due to *a* in Example 1.

Table 2. Varying a for the solutions of Example 1.

a	$J(u^*)$	$J(\bar{u})$
-0.10	31.311	20.407
-0.05	27.548	18.684
0.03	24.059	18.195
0.08	28.892	19.313

where

$$e(x) = e^{\left[\frac{25x}{2+x}\right]}.$$

The initial conditions are:

$$x_1(0) = 0.1 + a$$

$$x_2(0) = 0. \tag{55b}$$

The objective functional is:

$$J = \frac{1}{2}\int_0^1 (100(x_1(t))^2 + 100(x_2(t))^2 + (u(t))^2)dt. \tag{56}$$

Again, we have one coefficient a which occurs in the initial conditions and we set a = 0. As before, two problems are solved, one with minimizing the sensitivity factor and one without. Minimizing the sensitivity factor in this case means we want to minimize the "objective functional":

$$G = J + \alpha\left(\frac{\partial J}{\partial a}\right)^2, \tag{57}$$

where J is given by equation (56), subject to equation (55). This time n_p is taken as 40 and the weighting factor α is set to 0.01. Let u^* and $\bar{u}$ be the optimal controls of the problems (55) with equation (56) and (55) with equation (57), respectively. A plot of these is shown in Figure 3. Note that unlike Example 1, u^* and $\bar{u}$ are not very different. Let $J(u^*)$ and $J(\bar{u})$ be the values of J corresponding to u^* and $\bar{u}$, respectively. We observe that $\bar{u}$ manages to reduce the size of $\left\|\frac{\partial J}{\partial a}\right\|$, although $J(\bar{u}) = 0.74234 > J(u^*) = 0.73043$. See Table 3.

Recall that u^* and $\bar{u}$ are obtained with $a = 0$. We take them to be the base controls and proceed to vary the value of a. The results obtained are plotted in Figure 4. Although for negative values of a $J(\bar{u})$ is slightly greater than $J(u^*)$, there are some significant gains for positive values of a. The results for some selected values of a are shown in Table 4.

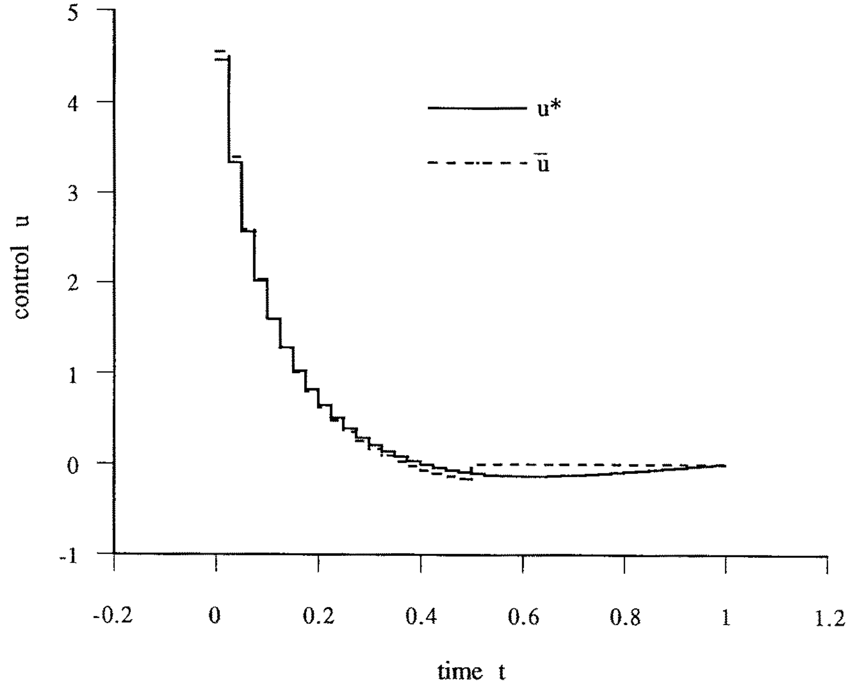

Figure 3. Optimal controls computed for Example 2.

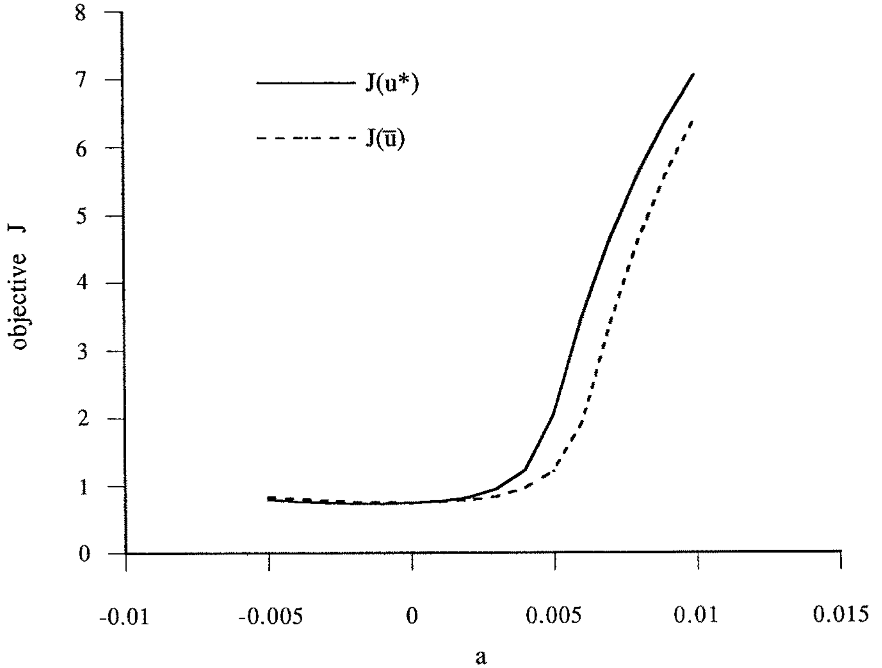

Figure 4. Variation in objective functional due to *a* in Example 2.

Table 3. Optimal solution for Example 2.

Problem	Objective J	$\left\| \frac{\partial J}{\partial a} \right\|$
(55) with (56)	0.73043	15.029
(55) with (57)	0.74234	0.00533

Table 4. Varying a for the solutions of Example 2.

a	$J(u^*)$	$J(\bar{u})$
-0.005	0.78511	0.82886
-0.002	0.72904	0.76099
0.003	0.94100	0.83393
0.005	2.0325	1.2118
0.006	3.4421	1.9135

5. Conclusions

In this paper, we consider a class of optimal control problems in which some of the coefficients are not known exactly. The aim is to find a control which minimizes the objective functional, and at the same time makes it less sensitive to the changes in the values of these coefficients. To solve this problem, a new problem is formulated so that the new objective functional consists of the original objective functional and a term which measures the sensitivity of the original objective functional with respect to the changes in the values of these uncertain coefficients. The gradient formula for the new objective functional is derived. On this basis, a computational method is proposed. For illustration, two examples are solved using the proposed method.

References

1. C.J. Goh and K.L. Teo, "Control Parameterization: a Unified Approach to Optimal Control Problems with General Constraints," *Automatica*, 24, pp. 3–18, 1988.
2. L. Hasdorff, *Gradient Optimization and Nonlinear Control*, John Wiley & Sons: New York, 1976.
3. L.S., Jennings, M.E. Fisher, K.L. Teo and C.J. Goh, "MISER3: Solving Optimal Control Problems — an Update," *Advances in Engineering Software and Workstations*, vol. 13, no. 4, pp. 190–196, 1991.
4. L. Lapidus, and R. Luus, *The control of nonlinear systems; part II: Convergence by combined first and second variations*, A.I. Ch. E.J., 13, pp. 108–113, 1967.
5. A. Miele, "Gradient Algorithms for the Optimization of Dynamic Systems," In C.T. Leondes (ed.), *Control and Dynamic Systems: Advances in Theory and Applications*, Academic Press: New York, 16, pp. 1–52, 1980.
6. N.B. Nedeljkovic, "New Algorithms for Unconstrained Nonlinear Optimal Control Problems," *IEEE Trans. Auto. Control*, AC-26, pp. 868-884, 1981.
7. K.L. Teo, C.J. Goh, and K.H. Wong, *A Unified Computational Approach for Optimal Control Problems*. Longman Scientific and Technical, England, 1991.