J. Korean Math. Soc. **37** (2000), No. 4, pp. 625–643

NECESSARY AND SUFFICIENT OPTIMALITY CONDITIONS FOR CONTROL SYSTEMS DESCRIBED BY INTEGRAL EQUATIONS WITH DELAY

Gamal N. Elnagar, Mohammad A. Kazemi, and Hoonjoo Kim

Abstract. In this paper we formulate an optimal control problem governed by time-delay Volterra integral equations; the problem includes control constraints as well as terminal equality and inequality constraints on the terminal state variables. First, using a special type of state and control variations, we represent a relatively simple and self-contained method for deriving new necessary conditions in the form of Pontryagin minimum principle. We show that these results immediately yield classical Pontryagin necessary conditions for control processes governed by ordinary differential equations (with or without delay). Next, imposing suitable convexity conditions on the functions involved, we derive Mangasarian-type and Arrow-type sufficient optimality conditions.

1. Introduction

Integral equations are a well-known mathematical tool for representing physical problems. Historically, they have achieved great popularity among mathematicians and physicists in formulating boundary-value problems of gravitation, electrostatics, fluid dynamics and scattering. Another application of considerable current interest is semiconductor modelling. It is also well known that initial-value and boundary-value problems for differential equations can often be converted into

Received March 3, 2000.

2000 Mathematics Subject Classification: 49J22.

Key words and phrases: optimal control, time-delay, Volterra integral equation, necessary condition, Pontryagin minimum principle, Mangasarian-type, Arrow-type, sufficient optimality condition.

The work of this author was supported in part by Deabul University and Korea Science and Engineering Foundation, 1997.

integral equations and there are usually significant advantages to be gained from making use of this conversion.

Optimization of control processes governed by differential equations with deviation arguments has been studied rather extensively in the literature [4, 9, 10, 12, 14, 19, 20]. With regard to optimal-control problems described by integral equations, however, it appears that only delay-free systems have been studied [1, 6, 11, 16, 17] and that only necessary conditions have been obtained. The aim of the present paper is to establish new necessary and sufficient conditions for optimality in a control process whose state is governed by a nonlinear Volterra integral equation with a nonlinear time delay in the state variable. Our formulation allows for restrictions on the control variables as well as the terminal state variables. First, following the theme developed in the works of [7, 13] for optimal control of differential equations, we develop a special set of state and control variations and show that this set forms a convex cone. Thus applying separation theorems of functional analysis, we obtain the basic necessary conditions. Next, we introduce an adjoint variable which is the solution to a linear Volterra integral equation, and derive a set of Pontryagin-type necessary conditions. These conditions are obtained under a rather general set of assumptions. In particular, unlike the results of [3, 11], we do not require that the state variable $x(t)$ be differentiable. The necessary conditions obtained here, of course, generalized those of [6]. We also show that these necessary conditions lead naturally to the well known optimality conditions for systems described by ordinary differential equations (with or without delay). Furthermore, motivated by [5], we introduce a suitable Hamiltonian function and obtain sufficient optimality conditions similar to those of Mangasarian [15] and Arrow[2] developed for optimal control problems described by ordinary differential equations. The sufficient conditions developed here have no counterpart in the literature.

2. Problem formulation

Let $I = [t_1, t_2]$ be a fixed closed and bounded interval in $\mathbb{R}$ and let $\sigma(t)$ be a strictly increasing absolutely continuous function satisfying $\sigma(t) \leq t$ and with inverse $\gamma(t)$. Let $S \subset \mathbb{R}^\nu$ and denote by $M(I, S)$ and $C(I, S)$, respectively, the set of all functions $F : I \to S$ which are measurable or continuous. Let m and n be positive integers, M and N be nonnegative integers, $X \subset \mathbb{R}^n$ be an open set, $U \subset \mathbb{R}^m$ be a convex open set and

$$\begin{aligned} \mathcal{X} &:= C(I, X), & \mathcal{U} &:= M(I, U), \\ g &: I \to \mathbb{R}^n, & h &: I \times I \times X \to \mathbb{R}^n, \\ k &: I \times I \to \mathbb{R}, & f &: I \times X \times X \times U \to \mathbb{R}^n, \\ \phi_i &: X \to \mathbb{R}, & i &= 0, 1, \cdots, M, \\ \psi_j &: X \to \mathbb{R}, & j &= M+1, \cdots, M+N. \end{aligned}$$

We are concerned with the following optimal control problem: Find $(x, u) \in \mathcal{X} \times \mathcal{U}$ that minimizes

$$J(x, u) := \phi_0\big(x(t_1)\big) \tag{2.1}$$

subject to
(2.2)

$$x(t) = g(t) + \int_{t_1}^{t} [h(t, s, x(s)) + k(t, s) f(s, x(s), x(\sigma(s)), u(s))] ds, \quad t \in I,$$

$$x(t) = \varphi(t), \quad t \in [\sigma(t_1), t_1]$$

$$\phi_i(x(t_2)) \leq 0, \quad i = 1, \cdots, M, \tag{2.3}$$

$$\psi_j(x(t_2)) = 0, \quad j = M+1, \cdots, M+N. \tag{2.4}$$

We assume that the following conditions hold:

(A1) g is continuous.
(A2) h is continuous with continuous first order partial derivatives w.r.t. $x \in \mathbb{R}^n$.
(A3) k is integrable on $I \times I$.
(A4) $f(t, x, y, u)$ is continuous with first order partial derivatives w.r.t. x, y and u.
(A5) φ is continuous on $[\sigma(t_1), t_1]$.
(A6) There is a nonnegative integrable function $\zeta(t)$ such that for all $(t, x, y, u) \in I \times X \times X \times U$, $|f| + |f_x| + |f_y| + |f_u| \leq \zeta(t)$.

In the notation above, f_x denotes the matrix of first partial derivatives of f w.r.t. x, that is,

$$f_x(t,x,y,u) := \left(\frac{\partial f_i}{\partial x_j}\right), \quad i=1,\dots,n, \quad j=1,\dots,n.$$

Similarly, for f_y and f_u.

(A7) ϕ_i and ψ_j are continuously differentiable on X.

Throughout this paper, $(\bar{x},\bar{u})$ will designate a solution to the optimal control (2.1)–(2.4), i.e., an optimal pair. Since many functions must be evaluated at $(\bar{x}(t),\bar{u}(t))$, this evaluation is represented compactly by a superbar. For example, we will write

$$\bar{f}(t) := f(t,\bar{x}(t),\bar{x}(\sigma(t)),\bar{u}(t))$$
$$\bar{f}_x(t) := f_x(t,\bar{x}(t),\bar{x}(\sigma(t)),\bar{u}(t)).$$

Since $\gamma(t)$ is the inverse of $\sigma(t)$, we write

$$\bar{f}_y\big(\gamma(t)\big) := f_y\Big(\gamma(t),\bar{x}\big(\gamma(t)\big),\bar{x}(t),\bar{u}(\gamma(t))\Big).$$

3. Cone of variations

We require a suitable set of state and control variations. The state variations are elements of $\tilde{\mathcal{X}} := C(I,\mathbb{R}^n)$. The set of control variations is any set $\tilde{\mathcal{U}}$ which has the following properties

$$\tilde{\mathcal{U}} \subset L_\infty(I,\mathbb{R}^m) \tag{3.1}$$

$$\tilde{\mathcal{U}} \subset \mathcal{U}-\bar{u} := \{w|w(t)=u(t)-\bar{u}(t), \quad u\in\mathcal{U}\} \tag{3.2}$$

$$0\in\tilde{\mathcal{U}}. \tag{3.3}$$

Now let $\epsilon=(\epsilon_1,\cdots,\epsilon_r)$ be a small parameter with $|\epsilon_j|<1$, $j=1,\cdots,r$, and $c_j\geq 0$, $j=1,\cdots,r$, and $\omega_j\in\mathcal{U}$, $j=1,\cdots,r$. Consider the perturbed integral equation

(3.4)

$$x_\epsilon(t)=g(t) + \int_{t_1}^{t}\Big[h(t,s,x_\epsilon(s))+k(t,s)f(s,x_\epsilon(s),x_\epsilon(\sigma(s)),\bar{u}(s)+\sum_{j=1}^{r}c_j\epsilon_j\omega_j(s))\Big]ds,$$

$$x_\epsilon(t)=\varphi(t), t\in[\sigma(t_1),t_1].$$

Notice that for sufficiently small ϵ, $\epsilon \geq 0$, $\bar{u} + \sum_{j=1}^{r} c_j \epsilon_j \omega_j$ takes values in the set U, and that for $\epsilon = 0$, $x_\epsilon = \bar{x}$ by the uniqueness of the solution of the integral equation. It also follows from our assumptions on h and f that $x_\epsilon(t)$ is differentiable w.r.t. ϵ.

Differentiating (3.4) w.r.t. ϵ_j and setting

$$W_j(t) := \Big(\frac{\partial x_\epsilon(t)}{\partial \epsilon_j}\Big)_{\epsilon=0}.$$

We derive the variational equation

$$\begin{aligned} W_j(t) = \int_{t_1}^{t} \Big\{ \Big[\bar{h}_x(t,s) + k(t,s)\bar{f}_x(s)\Big] W_j(s) \\ + k(t,s)\bar{f}_y(s) W_j(\sigma(s)) + c_j k(t,s)\bar{f}_u(s)\omega_j(s)\Big\}\, ds, \\ W_j(t) = 0, \quad t \in [\sigma(t_1), t_1]. \end{aligned} \tag{3.5}$$

Next we define a set K of variations as follows: Let $c \geq 0$, $\omega(t) \in \tilde{\mathcal{U}}$, and let $W(t)$ be the corresponding state variation satisfying (3.5). We denote this variation by $\nu := (c, \omega, W)$, and define

$$Y_k(\nu) := \Phi_x\big(\bar{x}(t_2)\big) W(t_2), \qquad k = 0, 1, \cdots, M+N,$$

where Φ denotes any of the ϕ_0, $\phi_1, \cdots, \phi_M$, $\psi_{M+1}, \cdots, \psi_{M+N}$. Define

$$K := \Big\{Y(\nu) := (Y_0, Y_1, \cdots, Y_{M+N}) \in \mathbb{R}^{1+M+N} | \nu = (c, \omega, W) \quad \text{a variation}\Big\}.$$

LEMMA 3.1. *The set K is a convex cone with vertex at the origin, that is, if $Y(\nu_1), Y(\nu_2) \in K$ and $a_1 \geq 0, a_2 \geq 0$, then $a_1 Y(\nu_1) + a_2 Y(\nu_2) \in K$.*

Proof. Let $\nu_j = (c_j, \omega_j, W_j)$, $j = 1, 2$, be two variations and $Y(\nu_j)$ the corresponding elements of K. Suppose first that $c := a_1 c_1 + a_2 c_2 \neq 0$, i.e., $c > 0$, and define the variation

$$\nu := \Big(a_1 c_1 + a_2 c_2, \frac{a_1 c_1 \omega_1 + a_2 c_2 \omega_2}{a_1 c_1 + a_2 c_2}, a_1 W_1 + a_2 W_2\Big).$$

Then it follows that $a_1Y(\nu_1) + a_2Y(\nu_2) = Y(\nu)$, an element of K. If $a_1c_1 + a_2c_2 = 0$, then $a_1Y(\nu_1) + a_2Y(\nu_2) = 0 = Y(\nu)$ corresponding to the variation $(0, 0, 0)$. This proves the lemma. □

In the following lemma, we require the following notation:

$$I_A := \{0\} \cup \Big\{ i | \phi_i(\bar{x}(t_2)) = 0, \ i = 1, \cdots, M \Big\}$$

i.e., I_A is the set of active inequalities.

LEMMA 3.2. *Let* $B := \Big\{ (b_0, b_1, \cdots, b_M, 0, 0, \cdots, 0) \in \mathbb{R}^{1+M+N} | b_i < 0$ *for* $i \in I_A \Big\}$. *Then* B *does not intersect the interior of the cone* K, *i.e.,*

$$B \cap \mathrm{Int}K = \emptyset.$$

Proof. Suppose on the contrary, that there is a point b with $b \in B \cap \mathrm{Int}K$. Thus $b = (b_0, b_1, \cdots, b_M, 0, 0, \cdots, 0)$, $b_i < 0$ for $i \in I_A$, and there is some $\delta > 0$ such that $1 + M + N$ points

$$\begin{array}{c}
(b_0,\ b_1 - \delta,\ b_2, \cdots,\ b_M,\ 0,\ 0,\ \cdots,\ 0) \\
(b_0,\ b_1,\ b_2 - \delta, \cdots,\ b_M,\ 0,\ 0,\ \cdots,\ 0) \\
\cdots \\
(b_0,\ b_1,\ \cdots, b_M - \delta,\ 0,\ 0,\ \cdots,\ 0) \\
(b_0,\ b_1,\ \cdots, b_M,\ -\delta,\ 0,\ 0,\ \cdots,\ 0) \\
(b_0,\ b_1,\ \cdots, b_M,\ 0,\ -\delta,\ 0,\ \cdots,\ 0) \\
\cdots \\
(b_0,\ b_1,\ \cdots, b_M,\ 0,\ 0,\ \cdots,\ 0,\ -\delta,\ 0) \\
(b_0,\ b_1,\ \cdots, b_M,\ 0,\ 0,\ \cdots,\ 0,\ -\delta) \\
(b_0,\ b_1,\ \cdots, b_M,\ \delta,\ \delta,\ \cdots,\ \delta)
\end{array} \tag{3.6}$$

all belong to K. Therefore, there are $1 + M + N$ variations $\nu_1, \cdots, \nu_{1+M+N}$ such that the corresponding vectors $Y(\nu_1), \cdots, Y(\nu_{1+M+N})$ are exactly those in (3.6).

We choose $r = 1 + M + N$ in (3.4), and consider the nonlinear algebraic system of $1 + M + N$ equations

$$(3.7)\qquad \begin{aligned} &\phi_i\big(x_\epsilon(t_2)\big) + a_i\epsilon_0 - \phi_i\big(\bar{x}(t_2)\big) = 0, && i = 0, 1, \cdots, M \\ &\psi_j\big(x_\epsilon(t_2)\big) - \psi_j\big(\bar{x}(t_2)\big) = 0, && j = M+1, \cdots, M+N \end{aligned}$$

where ϵ_0 is unknown, $a_0 = 1$, $a_i := b_0^{-1}\big[b_i - (1 + M + N)^{-1}\delta\big]$ for $i \in I_A - \{0\}$; $a_i = 0$ for $i \notin I_A$. (3.7) is a system of $1 + M + N$ equations in $2 + M + N$ unknowns $\epsilon_0, \epsilon = (\epsilon_1, \cdots, \epsilon_{1+M+N})$. These equations are obviously satisfied by $\epsilon_0 = 0, \epsilon = 0$, by the uniqueness of the solution of the integrable equation. At $\epsilon = 0$, the partial derivative with respect to ϵ_k of the left-hand side of (3.7) yield

$$(3.8)\qquad \begin{aligned} &\bar{\phi}_{ix}\big(\bar{x}(t_1)\big)W_k(t_2) = Y_i(\nu_k), && i = 0, 1, \cdots, M \\ &\bar{\psi}_{jx}\big(\bar{x}(t_1)\big)W_k(t_2) = Y_j(\nu_k), && j = 1 + M, \cdots, M + N. \end{aligned}$$

Thus according to (3.8) the Jacobian of (3.7) which is just the determinant of the $(1 + M + N) \times (1 + M + N)$ matrix whose columns are given by vectors in (3.6), is nonzero. It therefore follows from the implicit function theorem that for $\epsilon_0 \neq 0$ and sufficiently small, the system of equations (3.7) can be solved for $\epsilon_1, \cdots, \epsilon_{1+M+N}$ in terms of ϵ_0 and that for such an ϵ_0, the solutions

$$\epsilon_j := e_j(\epsilon_0), \qquad j = 1, \cdots, 1 + M + N$$

are continuously differentiable functions of ϵ_0, i.e., there is $\lambda > 0$ such that for $|\epsilon_0| < \lambda$, the functions $e(\epsilon_0) = \big(e_1(\epsilon_0), \cdots, e_{1+M+N}(\epsilon_0)\big)$ satisfy (3.7). At $\epsilon_0 = 0$, the partial derivatives of (3.7) with respect to ϵ_0 yield

$$(3.9)\qquad \begin{aligned} &\sum_{k=1}^{1+M+N} Y_i(\nu_k)e_k{}'(0) = -a_i, && i = 0, 1, \cdots, M \\ &\sum_{k=1}^{1+M+N} Y_j(\nu_k)e_k{}'(0) = 0, && j = M+1, \cdots, M+N \end{aligned}$$

where the columns of the matrix of coefficients $\big(Y_r(\nu_k)\big)$, $r = 0, 1, \cdots, M + N$, $k = 1, \cdots, 1 + M + N$ are given by vectors (3.6). By direct calculations we get, from (3.9),

$$e_j{}'(0) = -b_0^{-1}[1 + M + N]^{-1}, \qquad j = 1, \cdots, 1 + M + N.$$

Notice that $e_j'(0) > 0$ and that is the only solution of (3.9). We therefore conclude that for $\epsilon_0 > 0$ and sufficiently small, say $0 < \epsilon_0 \leq \lambda$, the numbers $\epsilon_j = e_j(\epsilon_0)$ are all positive and as close to zero as we wish (since $e_j(0) = 0$). Thus for $0 < \epsilon_0 \leq \lambda$, it follows from (3.7) that the pair $(x_\epsilon(t), u_\epsilon(t))$ with $\epsilon = e(\epsilon_0)$ and $u_\epsilon(t) := \bar{u}(t) + \sum_{j=1}^{1+M+N} c_j \epsilon_j \omega_j(t)$, satisfies

$$\phi_0(x_\epsilon(t_2)) = -\epsilon_0 + \phi_0(\bar{x}(t_2)) < \phi_0(\bar{x}(t_2)), \tag{3.10}$$

$$\phi_i(x_\epsilon(t_2)) = -a_i\epsilon_0 + \phi_i(\bar{x}(t_2)) \leq \phi_i(\bar{x}(t_2)), \tag{3.11}$$

$$\psi_j(x_\epsilon(t_2)) = \psi_j(\bar{x}(t_2)) = 0. \tag{3.12}$$

Relations (3.11)–(3.12) imply that (x_ϵ, u_ϵ) is an admissible pair. However (3.10) contradicts the optimality of $\bar{x}$, $\bar{u}$. This contradiction proves the lemma. □

THEOREM 3.1. *There is a nonzero vector* $\chi = (\chi_0, \chi_1, \cdots, \chi_{M+N}) \in \mathbb{R}^{1+M+N}$ *satisfying the following conditions:*

(i) $\chi_j \geq 0$ *for* $k \in I_A$

(ii) $\chi_j = 0$ *for* $k \in \{0, 1, \cdots, M\} - I_A$

(iii) $\sum_{j=0}^{M+N} \chi_j Y_j(\nu) \geq 0$ *for all variations* $\nu = (c, \omega, W)$.

Proof. If K has no interior points, the theorem follows immediately. If the interior of K is not empty, by Lemmas 3.1 and 3.2 there is a supporting hyperplane through $(0, 0, \cdots, 0)$ defined by a nonzero vector $\chi = (\chi_0, \chi_1, \cdots, \chi_{M+N})$, say, $\sum_{j=0}^{M+N} \chi_j Y_j \geq 0$ such that K is contained in $\sum_{j=0}^{M+N} \chi_j Y_j \geq 0$ and the set B defined in Lemma 3.2 contained in the other side $\sum_{j=0}^{M+N} \chi_j Y_j \leq 0$. It then easily follows from the definition of B that χ must satisfy (i) and (ii). The proof is complete. □

4. Necessary conditions for optimality

In this section, we derive the desired necessary conditions for optimality based on Theorem 3.1. We begin with the following developments.

Consider the linear integral equation

(4.1)
$$W(t) = \int_{t_1}^{t} k(t,s)\bar{f}_u(s)\omega(s)\,ds + \int_{t_1}^{t} \big\{\big[\bar{h}_x(t,s) + k(t,s)\bar{f}_x(s)\big]W(s) + k(t,s)\bar{f}_y(s)W(\sigma(s))\big\}\,ds.$$

Through a change of variable in the term involving $W(\sigma(s))$ we rewrite the above equation as

$$W(t) = \int_{t_1}^{t} k(t,s)\bar{f}_u(s)\omega(s)\,ds + \int_{t_1}^{t} \big[\bar{h}_x(t,s) + k(t,s)\bar{f}_x(s) + \chi(s)\dot{\gamma}(s)k(t,\gamma(s))\bar{f}_y(\gamma(s))\big]W(s)\,ds,$$

where $\chi(s)$ is the characteristic function on the interval $[t_1, \sigma(t)]$ and $\gamma(s)$ is the inverse of $\sigma(s)$. Thus $W(t)$ can be represented as

$$W(t) = \int_{t_1}^{t} k(t,s)\bar{f}_u(s)\omega(s)ds + \int_{t_1}^{t} R(t,s)\big[\int_{t_1}^{s} k(s,\sigma)\bar{f}_u(\sigma)\omega(\sigma)d\sigma\big]\,ds, \tag{4.2}$$

where $R(t,s)$, $t_1 \le s \le t \le t_2$, is the resolvent kernel associated with the integral equation (4.1):

$$\begin{aligned} R(t,s) = {} & \bar{h}_x(t,s) + k(t,s)\bar{f}_x(s) + \chi(s)\dot{\gamma}(s)k(t,\gamma(s))\bar{f}_y(\gamma(s)) \\ & + \int_{s}^{t} R(t,\tau)\big[\bar{h}_x(\tau,s) + k(\tau,s)\bar{f}_x(s) + \chi(s)\dot{\gamma}(s)k(\tau,\gamma(s))\bar{f}_y(\gamma(s))\big]\,d\tau, \\ & t_1 \le s \le \sigma(t). \end{aligned}$$

We may rewrite this equation as

$$\begin{aligned} R(t,s) = {} & \bar{h}_x(t,s) + k(t,s)\bar{f}_x(s) + \dot{\gamma}(s)k(t,\gamma(s))\bar{f}_y(\gamma(s)) \\ & + \int_{s}^{t} R(t,\tau)\big[\bar{h}_x(\tau,s) + k(\tau,s)\bar{f}_x(s) + \dot{\gamma}(s)k(\tau,\gamma(s))\bar{f}_y(\gamma(s))\big]\,d\tau, \\ & t_1 \le s \le \sigma(t), \end{aligned} \tag{4.3}$$

(4.4)
$$R(t,s) = \bar{h}_x(t,s) + k(t,s)\bar{f}_x(s) + \int_{s}^{t} R(t,\tau)[\bar{h}_x(\tau,s) + k(\tau,s)\bar{f}_x(s)]d\tau, \quad \sigma(t) \le s \le t.$$

We now transform Theorem 3.1 into a Pontryagin-type minimum principle. In the following theorem, $L(\chi)$ denotes

$$L(\chi) := \sum_{i=0}^{M} \chi_i \phi_{ix}\Big(\bar{x}(t_2)\Big) + \sum_{j=M+1}^{M+N} \chi_j \psi_{jx}\Big(\bar{x}(t_2)\Big)$$

and $I_A := \{0\} \cup \Big\{k | \phi_k(\bar{x}(t_2))) = 0,\ k = 1, \cdots, M\Big\}$.

THEOREM 4.1. *Consider the optimal control problem (2.1)–(2.4) and assume that (A1)–(A7) hold. If $\big(\bar{x}(t), \bar{u}(t)\big)$ be an optimal pair, then there exists a nonzero vector $\chi = (\chi_0, \chi_1, \cdots, \chi_{M+N}) \in \mathbb{R}^{1+M+N}$ and an n-vector function $p : I \to \mathbb{R}^n$ satisfying the following conditions:*

(i) $\chi_k \geq 0$ *for* $k \in I_A$
(ii) $\chi_k = 0$ *for* $k \in \{0, 1, \cdots, M\} - I_A$
(iii)

$$\begin{aligned} p(t) = L(\chi)\Big[\bar{h}_x(t_2, t) + k(t_2, t)\bar{f}_x(t) + \dot{\gamma}(t)k(t_2, \gamma(t))\bar{f}_y(\gamma(t))\Big] \\ + \int_t^{t_2} p(s)\Big[\bar{h}_x(s, t) + k(s, t)\bar{f}_x(t) + \dot{\gamma}(t)k(s, \gamma(t))\bar{f}_y(\gamma(t))\Big] ds, \end{aligned}$$

for $t_1 \leq t \leq \sigma(t_2)$

$$\begin{aligned} p(t) = L(\chi)\Big[\bar{h}_x(t_2, t) + k(t_2, t)\bar{f}_x(t)\Big] \\ + \int_t^{t_2} p(s)\big[\bar{h}_x(s, t) + k(s, t)\bar{f}_x(t)\big] ds, \end{aligned}$$

for $\sigma(t_2) \leq t \leq t_2$

(iv) *for almost all* $t \in I$ *and all* $u \in \mathcal{U}$

$$\Big[k(t_2, t)L(\chi) + \int_t^{t_2} k(s, t)p(s)\, ds\Big] \bar{f}_u(t)\big(u(t) - \bar{u}(t)\big) \geq 0.$$

Proof. Take an arbitrary variation $\nu = (1, \omega, W)$ and apply Theorem 3.1 to get a nonzero vector $\chi \in \mathbb{R}^{1+M+N}$ satisfying (i) and (ii) above, and

$$L(\chi)W(t_2) \geq 0.$$

Substituting for $W(t_2)$ from (4.2), this inequality takes the form of

$$\begin{aligned}&\int_{t_1}^{t_2} L(\chi)k(t_2,s)\bar{f}_u(s)\omega(s)ds\\ (4.5)\qquad &+\int_{t_1}^{t_2} L(\chi)R(t_2,s)\Big[\int_{t_1}^{s} k(s,\sigma)\bar{f}_u(\sigma)\omega(\sigma)d\sigma\Big]ds \geq 0.\end{aligned}$$

Next define the function $p : I \to \mathbb{R}^n$ by

$$p(t) = L(\chi)R(t_2, t).$$

By the definition of R, it is easily seen that $p(t)$ satisfies (iii) and that (4.5) takes the form of (iv). Thus the theorem follows. □

An application of Theorem 4.1

Theorem 4.1 generalizes the classical Pontryagin necessary conditions for control systems governed by ordinary differential equations with delay. Indeed, if $g(t) = x_0 = \varphi(t)$ (a constant), $h(t,s,x) \equiv 0$, and $k(t,s) \equiv 1$, then the integral equation (2.2) becomes equivalent to the ordinary differential equation $\dot{x}(t) = f\big(t, x(t), x(\sigma(t)), u(t)\big)$ with $x(\theta) = x_0$, for $\theta \in [\sigma(t_1), t_1]$. Defining the adjoint function

$$\lambda(t) := L(\chi) + \int_t^{t_2} p(s)ds \tag{4.6}$$

and the Hamiltonian H by

$$H(t,x,y,u,\lambda) := \lambda f(t,x,y,u),$$

we have, from (iii) and (4.6),

$$\begin{aligned}&\dot{\lambda}(t) = -\bar{H}_x(t) - \dot{\gamma}(t)H_y\big(\gamma(t)\big), \quad \text{for } \ t_1 \leq t \leq \sigma(t_2),\\ &\dot{\lambda}(t) = -\bar{H}_x(t), \quad \text{for } \ \sigma(t_2) \leq t \leq t_2,\\ &\lambda(t_2) = L(\chi)\end{aligned}$$

and (iv) takes the form of

$$\bar{H}_u(t)\big(u(t) - \bar{u}(t)\big) \geq 0$$

for all admissible control functions u and for a.a. $t \in [t_1, t_2]$.

5. Lagrange-type objective function

In this section, we assume that the objective function (2.1) is given by

$$J(x,u) := \int_{t_1}^{t_2} K\Big(t, x(t), x(\sigma(t)), u(t)\Big) dt. \tag{5.1}$$

We require the following assumption regarding $K(t,x,y,u)$:

(A8) $K : I \times \mathbb{R}^n \times \mathbb{R}^n \times U \to \mathbb{R}$ is continuous, together with its first partial derivatives w.r.t. x, y and u. Furthermore, there is a nonnegative integrable function $\zeta(t)$ for all $(t,x,y,u) \in I \times X \times X \times U$, $|K| + |K_x| + |K_y| + |K_u| \leq \zeta(t)$.

We now derive the corresponding necessary conditions similar to those of Theorem 4.1. To this end, we introduce the auxiliary state variable

$$x^{n+1}(t) := \int_{t_1}^{t} K\Big(s, x(s), x(\sigma(s)), u(s)\Big)\, ds, \tag{5.2}$$

define $\tilde{x}(t) := \Big(x(t), x^{n+1}(t)\Big) \in \mathbb{R}^{n+1}$ and rewrite $J(x,u)$ as

$$J(\tilde{x}, u) := x^{n+1}(t_2). \tag{5.3}$$

Thus the optimal control problem (5.1) with (2.2)–(2.4) is reformulated as special case of problem (2.1)–(2.4) with the new state vector $\tilde{x}(t)$ satisfying

$$\tilde{x}(t) = \tilde{g}(t) + \int_{t_1}^{t} \Big[\tilde{h}\big(t, s, \tilde{x}(s)\big) + \tilde{k}(t,s)\tilde{f}\big(s, \tilde{x}(s), \tilde{x}(\sigma(s)), u(s)\big)\Big]\, ds,$$

where

$$\tilde{f} = \begin{pmatrix} f \\ K \end{pmatrix}, \quad \tilde{g} = \begin{pmatrix} g \\ 0 \end{pmatrix}, \quad \tilde{h} = \begin{pmatrix} h \\ 0 \end{pmatrix}.$$

In this case, we may think of $\bar{k}$ as the matrix $\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$. We apply Theorem 4.1 to problem (5.3), (2.2)–(2.4): In the present case,

$$\tilde{L}(\chi) = \big(L(\chi), \chi_0\big),$$
$$L(\chi) = \sum_{i=0}^{M} \chi_i \phi_{ix}\big(\bar{x}(t_2)\big) + \sum_{j=M+1}^{M+N} \chi_j \psi_{jx}\big(\bar{x}(t_2)\big),$$
$$\tilde{p}(t) = \big(p(t), p^{n+1}(t)\big).$$

Therefore, taking $p^{n+1}(t) = 0$, we see that, in this case,

(5.4)
$$\begin{aligned} p(t) = \chi_0 &\big[\bar{K}_x(t) + \dot{\gamma}(t)\bar{K}_y(\gamma(t))\big] \\ &+ L(\chi)\big[\bar{h}_x(t_2,t) + k(t_2,t)\bar{f}_x(t) + \dot{\gamma}(t)k(t_2,\gamma(t))\bar{f}_y(\gamma(t))\big] \\ &+ \int_t^{t_2} p(s)\big[\bar{h}_x(s,t) + k(s,t)\bar{f}_x(t) + \dot{\gamma}(t)k(s,t)\bar{f}_y(\gamma(t))\big]\, ds, \end{aligned}$$

for $t_1 \le t \le \sigma(t_2)$;

$$\begin{aligned} p(t) = \chi_0\bar{K}_x(t) &+ L(\chi)\big[\bar{h}_x(t_2,t) + k(t_2,t)\bar{f}_x(t)\big] \\ &+ \int_t^{t_2} p(s)\big[\bar{h}_x(s,t) + k(s,t)\bar{f}_x(t)\big]\, ds, \end{aligned} \tag{5.5}$$

for $\sigma(t_2) \le t \le t_2$.

The condition (iv) in Theorem 4.1 becomes

(5.6)
$$\big[\chi_0\bar{K}_u(t) + \big\{L(\chi)k(t_2,t) + \int_t^{t_2} k(s,t)p(s)\, ds\big\}\bar{f}_u(t)\big]\big(u(t) - \bar{u}(t)\big) \ge 0$$

for a.a. $t \in I$ and all $u \in \mathcal{U}$.

We summarize the preceding results in the following theorem:

THEOREM 5.1. *Consider the optimal control problem* (5.1) *with* (2.2)–(2.4). *And assume that* (A1)–(A8) *hold. If* $(\bar{x}, \bar{u})$ *is an optimal* *pair, then there exists a nonzero vector* $\chi = (\chi_0, \chi_1, \cdots, \chi_{M+N})$ *satisfying* (i) *and* (ii) *in* Theorem 4.1 *and a function* $p : [t_1, t_2] \to \mathbb{R}^n$ *such that* (5.4)–(5.6) *hold.*

6. Sufficient conditions

In this section, we establish sufficient optimality conditions for the optimal control problem (5.1) with (2.2)–(2.4). We define the Hamiltonian function H formally as

$$\begin{aligned} H(t,x,y,u,p) :=& \chi_0 K(t,x,y,u) + L(\chi)\big[h(t_2,t,x) + k(t_2,t)f(t,x,y,u)\big] \\ & + \int_t^{t_2} p(s)\big[h(s,t,x) + k(s,t)f(t,x,y,u)\big]ds \end{aligned} \tag{6.1}$$

where $L(\chi)$ is as defined in section 5. Then in terms of H, Theorem 5.1 states that if $(\bar{x}, \bar{u})$ is an optimal pair, then there exist multipliers χ_k, $0 \le k \le M+N$ and a function $p : I \to \mathbb{R}^n$ such that

$$(\chi_0, \chi_1, \cdots, \chi_{M+N}) \neq (0, 0, \cdots, 0); \quad \chi_k \ge 0 \quad \text{for} \quad 0 \le k \le M;$$

$$\chi_k \phi_k(\bar{x}(t_2)) = 0, \quad 1 \le k \le M; \tag{6.2}$$

$$p(t) = \bar{H}_x(t) + \dot{\gamma}(t)\bar{H}_y((\gamma(t)), \qquad t_1 \le t \le \sigma(t_2) \tag{6.3}$$

$$p(t) = \bar{H}_x(t), \qquad \sigma(t_2) \le t \le t_2 \tag{6.4}$$

$$\bar{H}_u(t)(u(t) - \bar{u}(t)) \ge 0 \quad \text{for a.a.} t \in I \quad \text{and all } u \in \mathcal{U}. \tag{6.5}$$

DEFINITION 6.1. A real-valued function F defined on a convex set C in $\mathbb{R}^n$ is said to be *quasi-convex* if

$$F(x) \le F(\bar{x}) \Rightarrow F(\theta x + (1-\theta)\bar{x}) \le F(\bar{x})$$

for all x, $\bar{x} \in C$ and all $\theta \in [0,1]$.

It can be easily shown that if F is real-valued, quasi-convex on a convex set C in $\mathbb{R}^n$, then

$$F(x) \le F(\bar{x}) \Rightarrow F'(\bar{x}) \cdot (x - \bar{x}) \le 0 \quad \text{for all} \quad x \in C.$$

THEOREM 6.1. *Let $(\bar{x}, \bar{u})$ be an admissible pair for problem (5.1) with (2.2)–(2.4). If there are multipliers $\chi_0, \chi_1, \cdots, \chi_{M+N}$ and a function $p(t)$ satisfying, for $\chi_0 = 1$, conditions (6.2)–(6.5) along with the following conditions*

$$\sum_{i=1}^{M} \chi_i \phi_i(x) + \sum_{j=M+1}^{M+N} \chi_j \psi_j(x) \text{ is quasi-convex in } x \tag{6.6}$$

$$H(t, x, y, u, p(t)) \text{ is convex in } (x, y, u) \text{ for a.a. } t \in I \tag{6.7}$$

then $(\bar{x}, \bar{u})$ is an optimal pair for problem (5.1) with (2.2)–(2.4).

Proof. Let (x, u) be any other admissible pair. We must prove that

$$\Delta J := \int_{t_1}^{t_2} [K(t, x(t), x(\sigma(t)), u(t)) - \bar{K}(t)]dt \geq 0. \tag{6.8}$$

With $\chi_0 = 1$ in (6.1), we can write

$$\begin{aligned}
&\Delta J \\
&= \int_{t_1}^{t_2} [H(t, x(t), x(\sigma(t)), u(t), p(t)) - \bar{H}(t)]dt \\
&\quad - L(\chi) \int_{t_1}^{t_2} \{h(t_2, t, x(t)) - \bar{h}(t_2, t) + k(t_2, t)[f(t, x(t), x(\sigma(t)), u(t)) - \bar{f}(t)]\}dt \\
&\quad - \int_{t_1}^{t_2} dt \int_{t}^{t_2} p(s)\{h(s, t, x(t)) - \bar{h}(s, t) + k(s, t)[f(t, x(t), x(\sigma(t)), u(t)) - \bar{f}(t)]\}ds \\
&\geq \int_{t_1}^{t_2} [\bar{H}_x(t) \cdot \Delta x(t) + \bar{H}_y(t)\Delta x(\sigma(t)) + \bar{H}_u(t) \cdot \Delta u(t)]dt \\
&\quad - L(\chi)\Delta x(t_2) - \int_{t_1}^{t_2} p(s)\Delta x(s)ds \\
&\geq \int_{t_1}^{\sigma(t_2)} [\bar{H}_x(t) + \dot{\gamma}(t)\bar{H}_y(\gamma(t))]\Delta x(t)dt + \int_{\sigma(t_2)}^{t_2} \bar{H}_x(t)\Delta x(t)dt \\
&\quad - L(\chi)\Delta x(t_2) - \int_{t_1}^{t_2} p(s)\Delta x(s)ds
\end{aligned}$$

where we have used (6.7), changed the order of integration, (6.5) and (2.2). Using (6.3) and (6.4), we simplify to get

(6.9)

$$\begin{aligned}
\Delta J &\geq -L(\chi)\Delta x(t_2) \\
&= -\{\sum_{i=1}^{M} \chi_i \phi_{ix}(\bar{x}(t_2)) + \sum_{j=M+1}^{M+N} \chi_j \psi_{jx}(\bar{x}(t_2))\}(x(t_2) - \bar{x}(t_2)).
\end{aligned}$$

Now let us denote

$$D(x) = \sum_{i=1}^{M} \chi_i \phi_i(x(t_2)) + \sum_{j=M+1}^{M+N} \chi_j \psi_j(x(t_2)).$$

Then

(6.10)
$$\begin{aligned}
&D(x(t_2)) - D(\bar{x}(t_2)) \\
&= \sum_{i=1}^{M} \chi_i(\phi_i(x(t_2)) - \phi_i(\bar{x}(t_2))) + \sum_{j=M+1}^{M+N} \chi_j(\psi_j(x(t_2)) - \psi_j(\bar{x}(t_2))) \\
&= \sum_{i=1}^{M} \chi_i(\phi_i(x(t_2)) - \phi_i(\bar{x}(t_2))) \\
&= \sum_{i=1}^{M} \chi_i \phi_i(x(t_2)) \leq 0
\end{aligned}$$

by (2.4), (6.2) and (2.3). Thus $D(x(t_2)) \leq D(\bar{x}(t_2))$. Finally, using (6.10) and the quasi-convex condition (6.6) for $D(x)$ we conclude that the right hand side of (6.9) is nonnegative. This completes our proof.□

REMARK 6.1. Theorem 6.1 generalizes the Mangasarian sufficient condition for optimal control process described by ordinary differential equations.

THEOREM 6.2. *Let $(\bar{x}, \bar{u})$ be an admissible pair for problem* (5.1) *with* (2.2)–(2.4). *If there exist multipliers $\chi_1, \cdots, \chi_{M+N}$ and a function $p(t)$ satisfying, for $\chi_0 = 1$, conditions* (6.2)–(6.4) *with* (6.6) *along with the following conditions*

(6.11) $\bar{H}(t) \leq H\big(t, \bar{x}(t), \bar{x}(\sigma(t)), v, p(t)\big)$, *for* a.a. $t \in [t_1, t_2]$ *and all* $v \in U$

(6.12) $\mathcal{H}\big(t, x, y, p(t)\big) := \min_{u \in U} H\big(t, x, y, u, p(t)\big)$ *exists and is convex in* (x, y) *for* a.a. t.

Then $(\bar{x}, \bar{u})$ is an optimal pair.

Proof. Suppose that (x,u) is an arbitrary admissible pair. We have to prove that ΔJ defined in (6.8) is nonnegative. Let us first note, from the definition of $\mathcal{H}$, that $\bar{H}(t) = \mathcal{H}(t,\bar{x}(t),\bar{x}(\sigma(t)),p(t))$ and that $H(t,x,y,u,p) \geq \mathcal{H}(t,x,y,p)$. Hence

(6.13)
$$H(t,x,y,u,p(t)) - \bar{H}(t) \geq \mathcal{H}(t,x,y,p) - \mathcal{H}(t,\bar{x}(t),\bar{x}(\sigma(t)),p(t)).$$

Now, from the convexity of $\mathcal{H}(t,x,y,p)$ it follows (cf. [18]) that $\mathcal{H}$ has a subgradient (a,b) at $\bar{x}$:

$$\mathcal{H}(t,x,y,p)-\mathcal{H}(t,\bar{x}(t),\bar{x}(\sigma(t)),p(t)) \geq \dot{a}\cdot(x-\bar{x}(t))+b\cdot(y-\bar{y}(t)) \tag{6.14}$$

for all $(x,y) \in \mathbb{R}^{2n}$. We next define, for fixed $t \in I$,

$$h(x,y) := H(t,x,y,u,p(t)) - \bar{H}(t) - a\cdot(x-\bar{x}(t)) - b\cdot(y-\bar{x}(\sigma(t))).$$

It follows from (6.13)–(6.14) that $h(x,y) \geq 0$ for all $(x,y) \in \mathbb{R}^{2n}$. Since $h(\bar{x}(t),\bar{x}(\sigma(t))) = 0$, $(\bar{x}(t),\bar{x}(\sigma(t)))$ minimizes h and so

$$\frac{\partial h}{\partial x}(\bar{x}(t),\bar{x}(\sigma(t))) = \frac{\partial h}{\partial y}(\bar{x}(t),\bar{x}(\sigma(t))) = 0,$$

i.e., $a = \bar{H}_x(t)$ and $b = \bar{H}_y(t)$. Therefore for all $(x,y,u) \in \mathbb{R}^n \times \mathbb{R}^n \times U$ and a.a. $t \in I$,

$$H(t,x,y,u,p(t)) - \bar{H}(t) \geq H_x(t)\cdot(x-\bar{x}(t)) + H_y(t)\cdot(y-\bar{x}(\sigma(t))).$$

The rest of the proof now proceeds as in the proof of Theorem 6.1. □

References

[1] Angell, T. S., *On the optimal control of systems governed by nonlinear integral equations*, J. Optimiz. Theory and Applics. **19**(1976), 63–79.

[2] Arrow, K. J. and Kurz, M., *Public Investment, the Rate of Return, and Optimal Fiscal Policy*, John Hopkins Press, Baltimore, 1970.

[3] Bakke, V. L., *A maximum principle for an optimal control problem with integral constraints*, J. Optimiz. Theory and Applics. **13** (1974) 32–55.

[4] Bank, H. T. and Jacobs, M , *An attainable sets approach to optimal control of functional differential equations with function space conditions*, J. Differential equations **13** (1973), 127–149.

[5] Burnap, C. and Kazemi, M. A., *Optimal control of a system governed by nonlinear Volterra integral equations with delay*, IMA J. Math. Control Info. **15** (1999), 1–17.

[6] Carlson, D. A., *An elementary proof of the maximum principle for optimal control problems governed by a Volterra integral equation*, J. Optimiz. Theory and Applics. **54** (1987), 43–61.

[7] Cesari, L., *Optimization-Theory and applications, problems with ordinary differential equations*, Springer-Verlag, New York, Hidelburg, Berlin, 1983.

[8] Corduneanu, C., *Integral equations and applications*, Cambridge University Press, Cambridge, 1991.

[9] El'sgol'c, L.E., *Qualitative methods of mathematical analysis*, Transl. AMS. American Mathematical Society, Providence, RI, 1964.

[10] Giorecki, H., *Analysis and synthesis of time delay systems*, Wiley, New York, 1989.

[11] Hartl, R. F. and Sethi, S. P., *Optimal control on a class of systems with continuous lags: a dynamic programming approach and economic interpretations*, J. Optimiz. Theory and Applics. **43** (1984), 73–88.

[12] Kazemi, M. A., *Necessary conditions for constrained distributed parameter systems with deviating argument*, J. Austral. Math. Soc. B **37** (1996), 495–511.

[13] ———, *Optimization of systems governed by hyperbolic partial differential equation with equality and inequality constraints*, Rendiconti Del Circolo Matematico Di Palermo Serie II, Tomo **34** (1985), 283–299.

[14] Kowalewski, A., *Optimal control of distributed parabolic systems involving time lags*, IMA J. Math. Control Info. **7** (1991), 375–393.

[15] Mangasarian, O. L., *Sufficient conditions for the optimal control of nonlinear systems*, SIAM J. Control **4** (1966), 139–152.

[16] Neustadt, L. W., *Optimization: a theory of necessary conditions*, Princeton University Press, Princeton, NJ, 1976.

[17] Pritchard, A. J and You, Y., *Causal feedback optimal control for Volterra integral equations*, SIAM J. Control and Optim. **34** (1996), 1874–1890.

[18] Rockafellar, R. T., *Convex analysis*, Princeton University Press, Princeton, NJ, 1970.

[19] Sadek, L. S., *Optimal control of time-delay systems with distributed parameters*. J. Optimiz. Theory and Applics. (1990), 567–585.

[20] Wong, P. K. C., *Optimal control of parabolic systems with boundary conditions having time delays*, SIAM J. Control and Optim. **13** (1975), 274–293.

Department of Mathematics
University of South Carolina at Spartanburg
Spartanburg, South Carolina 29303
USA
E-mail: mkazemi@email.uncc.edu

Department of Mathematics
University of North Carolina at Charlotte
Charlotte, North Carolina 28223
USA

Department of Computer Science
Daebul University
Young-Am 526-890
Korea
E-mail: hoonjoo@mail.daebul.ac.kr