Necessary conditions of optimality for a class of optimal control problems on time scales

Y. Peng\textsuperscript{a,b,*}, X. Xiang\textsuperscript{b}, Y. Gong\textsuperscript{b}, G. Liu\textsuperscript{b}

\textsuperscript{a} College of Computer Science and Information, Guizhou University, Guiyang, Guizhou 550025, PR China
\textsuperscript{b} Department of Mathematics, Guizhou University, Guiyang, Guizhou 550025, PR China

\textbf{ABSTRACT}

A class of optimal control problems of a system governed by linear dynamic equations on time scales with quadratic cost functional is considered. By the Lebesgue $\Delta$-integral theory and the Sobolev-type space $H^1$ on time scales the weak solution of linear dynamic equations on time scales for both initial value problem and backward problem are introduced, therefore the necessary conditions of optimality are presented. Some typical examples are given for demonstration.

1. Introduction

The calculus of time scales was initiated by Hilger in his Ph.D. thesis \cite{1} in 1988 in order to create a theory that can unify discrete and continuous analyses. The time scales calculus has a tremendous potential for applications in some mathematical models of real processes and phenomena studied in physics, chemical technology, population dynamics, biotechnology and economics, neural networks, social sciences, as is pointed out in the monographs of Hilger \cite{1}, Bohner \cite{2}, Brandt \cite{3} and Lakshmikantham et al. \cite{4}.

In recent years dynamic equations on time scales are considered for both initial value problems and boundary value problems. Some results on the existence, uniqueness and properties of classical solution were obtained. In addition, many authors including us studied the existence of classical solution for impulsive dynamic equations on time scales (see \cite{1–11}).

However, to our knowledge, optimal control problems of a system governed by dynamic equations on time scales have not been extensively developed. In 2006, we first considered a Lagrange problem (P) of a system governed by linear dynamic equations on time scales. Introducing the weak solution of linear dynamic equations we presented the existence of optimal controls for a class of optimal control problems (see \cite{12}).

It is well known that the Pontryagin maximum principle plays a central role in optimal control theory. In this paper, we present the necessary conditions of optimality for the Lagrange problem (P) of systems governed by linear dynamic equations on time scales with quadratic cost functional. We have not found any similar result in the literature.

In order to derive the necessary conditions of optimality containing optimal controlled system, adjoint equation and optimal inequality, we must study the linear dynamic equation on time scale for both the Cauchy problem and the backward problem. Since the properties of admission control functions are often more weak, classical solutions are not suitable for this problem. By the virtue of Lebesgue $\Delta$-integral and the Sobolev-type space $H^1$ on time scales we discuss the integral
functions with variable limit on time scales and their properties and introduce suitable weak solutions for both of the Cauchy problem and the backward problem. Utilizing the integration by parts in $H^1$ we derive the adjoint equation and the optimal inequality. The necessary conditions of optimality are presented at last. These results are helpful to discover relation and difference among the continuous optimal control problems, the discrete optimal control problems and the mathematical programming and establish uniform theory further.

The paper is organized as follows. In Section 2, we recall some basic definitions and facts from the time scales calculus that will be used in the sequel. Section 3 devotes the weak solution of Cauchy problem and backward problem for linear dynamic equations on time scales. Our main result is given in Section 4. The optimal control problem on time scales is formulated and corresponding necessary conditions of optimality are derived. Some typical examples are given for demonstration in the Section 5.

2. Preliminaries

In this section we briefly introduce the time scales calculus. A time scale $T$ is a nonempty closed subset of $R$. The two most popular examples are $T = R$ and $T = Z$. Define the forward and backward jump operators $\sigma, \rho : T \rightarrow T$ by

$$
\sigma(t) = \inf\{s \in T \mid s > t\}, \quad \rho(t) = \sup\{s \in T \mid s < t\}, \quad t \in T,
$$

where, in this definition, we write $\sup \emptyset = \inf T \equiv a$ and $\inf \emptyset = \sup T \equiv b$. A point $t \in T$ is said to be left-dense, left-scattered, right-dense, right-scattered if $\rho(t) = t$, $\rho(t) < t$, $\sigma(t) = t$, $\sigma(t) > t$, respectively. We assume that $\rho(b) > a$, so that $T$ must contain at least 3 points. Now define the set

$$
T^k = \begin{cases} T \setminus \{b\}, & \rho(b) < b, \\ T, & \rho(b) = b. \end{cases} \quad T_k = \begin{cases} T \setminus \{a\}, & \sigma(a) > a, \\ T, & \sigma(a) = a. \end{cases}
$$

The sets $T^k$ and $T_k$ are time scales. The forward (backward) graininess $\mu : T^k \rightarrow [0, +\infty)(\nu : T_k \rightarrow [0, +\infty))$ is defined by $\mu(t) = \sigma(t) - t(\nu(t) = t - \rho(t))$.

**Definition 2.1.** A function $f : T \rightarrow R$ is $\Delta$-differentiable at $t \in T^k$, if there exists a number $f^\Delta(t)$, with following property: for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$
|f(\sigma(t)) - f(s) - f^\Delta(t)(\sigma(t) - s)| \leq \varepsilon|\sigma(t) - s| \quad \text{for all } s \in T \text{ and } |s - t| < \delta.
$$

If $f$ is $\Delta$-differentiable for every $t \in T^k$, we say that $f$ is $\Delta$-differentiable on $T$.

Obviously, $1^\Delta = 0, t^\Delta = 1, (t^2)^\Delta = t + \sigma(t), (t^3)^\Delta = t^2 + t\sigma(t) + \sigma^2(t)$ and $(\frac{1}{t})^\Delta = -\frac{1}{t\sigma(t)}$. This means that if $T = R$ this definition coincides with classical derivative definition and if $T = Z$ it coincides with forward difference.

At the same time the Riemann $\Delta$-integral was introduced by partition idea. About the calculus and theory of time scales we refer the reader to the book of Bohner [2] and the papers of Cabada [13] and Guseinov [14]. A function $f : T \rightarrow R$ is said to be rd-continuous on $T$ if it is continuous at all right-dense points in $T$ and has finite left-sided limits at all left-dense points in $T$. Consider

$$
C_{id}(T, R) = \{f : T \rightarrow R \mid f \text{ is rd-continuous on } T\}
$$

endowed with norm

$$
\|f\|_{C_{id}} = \sup\{|f(t)| \mid t \in T\} \quad \text{for } f \in C_{id}(T, R).
$$

$C_{id}(T, R)$ is a Banach space.

Similarly, $C^1_{id}(T, R)$ is also a Banach space, where

$$
C^1_{id}(T, R) = \{f \in C_{id}(T, R), f^\Delta \in C_{id}(T^k, R)\}, \quad \|f\|_{C^1_{id}} = \|f\|_{C_{id}} + \|f^\Delta\|_{C_{id}} \quad \text{for } f \in C^1_{id}(T, R).
$$

A function $p : T \rightarrow R$ is called regressive if $1 + \mu(t)p(t) \neq 0, \forall t \in T$. The generalized exponential function $e_p$ is defined as the unique solution $y(t) = e_p(t, a)$ of the Cauchy problem $y^\Delta(t) = p(t)y(t), y(a) = 1$, where $p$ is a regressive and rd-continuous function. An explicit formula for $e_p(\cdot, \cdot)$ is given by

$$
e_p(t, s) = \exp \left\{ \int_s^t \xi_{\mu(t)}(p(t))\Delta t \right\} \quad \text{with } \xi_h(z) = \begin{cases} \frac{\ln(1 + zh)}{h} & \text{if } h \neq 0, \\ \frac{z}{2} & \text{if } h = 0. \end{cases}
$$

For more details, see [6,2]. Furthermore, $e_p(\cdot, \cdot)$ satisfies the semigroup properties

$$
e_p(t, s) = \begin{cases} 1, & e_p(t, 0) = e_p(s, 0) = e_p(t, s), \quad e_p(t, t) = 1. \end{cases}
$$

and the following fundamental operating properties

$$
e_p(\sigma(t), s) = (1 + \mu(t)p(t))e_p(t, s), \quad e_p(t, s) = \frac{1}{e_p(s, t)} = e_{\Delta p}(s, t),
$$

(2.1)
where
\[ \Theta p = -\frac{p}{1 + \mu p}. \]

Particularly, one can obtain the following analytic properties
\[ (e_p(\cdot, s))^\Delta = p(\cdot)e_p(\cdot, s), \quad (e_p(s, \cdot))^{\Delta} = -p(\cdot)e_p(s, \sigma(\cdot)). \tag{2.3} \]

The Lebesgue $\Delta$-measure $\mu_\Delta$ on time scale $T$ was defined as the Caratheodory extension of a set function in [15]. For each $t_0 \in T^k$ the single-point set $\{t_0\}$ is $\Delta$-measurable and its $\Delta$-measure is given by
\[ \mu_\Delta(\{t_0\}) = \sigma(t_0) - t_0. \]

It is clear that the Lebesgue $\Delta$-measure of single-point set on $T$ may not be equal to zero. In addition, The Lebesgue $\Delta$-measure $\mu_\Delta$ has closed association with Lebesgue measure $\mu$. $E$ is a subset of $T$. Define $\tilde{E} = E \bigcup_{i \in E} (t_i, \sigma(t_i))$, where $\sigma(t_i) - t_i > 0$. $\tilde{E}$ is called the extension of $E$. $E$ is Lebesgue $\Delta$-measurable if and only if $\tilde{E}$ is Lebesgue measurable and
\[ \mu_\Delta(E) = \mu(\tilde{E}) = \mu(E) + \sum_{i \in E} (\sigma(t_i) - t_i) = \mu(E) + \sum_{t_i \in E} \mu(t_i). \tag{2.4} \]

It is obvious that $\mu_\Delta(E) = \mu(E)$ if and only if $E$ has no right-scattered points.

As a straightforward consequence of equality (2.4), one can deduce the simple formula to calculate the Lebesgue $\Delta$-integral. For a function $f : T \rightarrow R$, define the step function interpolation $\tilde{f} : [a, b] \rightarrow R$ as
\[ \tilde{f}(t) = \begin{cases} f(t_i), & t \in (t_i, \sigma(t_i)), t_i \in T, \\ f(t) = f(t), & t \in T. \end{cases} \]

This function extends $f$ to the real interval $[a, b]$ and it allows us to establish an equivalence between Lebesgue $\Delta$-integral and classical Lebesgue integral. Let $E \subseteq T$ be $\Delta$-measurable set such that $b \notin E$. Now we say that $f$ is Lebesgue $\Delta$-integrable on $E$ if and only if $\tilde{f}$ is Lebesgue integrable on $\tilde{E}$, in which case the equality holds:
\[ \int_E f(s) \Delta s = \int_{\tilde{E}} \tilde{f}(s) ds. \]

The right side of the equality above is the classical Lebesgue integral over the real interval $\tilde{E} \subseteq [a, b]$. From this equality we obtain a important result which give us a formula for calculating the Lebesgue $\Delta$-integral. For all $s, t \in T$ with $s \leq t$, the following expression holds:
\[ \int_{[s, t]} f(\tau) \Delta \tau = \int_{[s, t] \cap T} f(\tau) d\tau + \sum_{s \leq h < t, t_i \in T} \mu(t_i) f(t_i). \tag{2.5} \]

Define
\[ L^1(T, R) = \{ f : T \rightarrow R \mid f \text{ is Lebesgue } \Delta \text{-integrable on } T \}, \quad L^p(T, R) = \{ f \in L^1(T, R) \mid \| f \|^p \in L^1(T, R) \}. \]

Endowed with norm
\[ \| f \|_{L^p} = \left( \int_{[a, b]} |f(\tau)|^p \Delta \tau \right)^{\frac{1}{p}}, \]

the spaces $L^p(T, R)$ ($p \geq 1$) is a Banach space. Clearly, $f \in L^1([a, b], R)$ means
\[ \int_a^t f(s) ds = 0 \quad \text{for any } t \in [a, b]. \]

But $f \in L^1(T, R)$ implies
\[ \int_{[s, t]} f(\tau) \Delta \tau = \mu(t) f(t) \quad \text{for any } t \in T \]
which may not be equal to zero.

In order study optimal control problem, we need Sobolev-type space and generalized derivative on time scales. Equipped with norm $\| \cdot \|_{H^2_\mu}$ on space $C^1_\mu$ given by
\[ \| f \|_{H^2_\mu} = \| f \|^2_{L^2} + \| f^\Delta \|^2_{L^2}, \quad f \in C^1_\mu(T, R). \]

The completion of $C^1_\mu(T, R)$ with respect to the norm $\| \cdot \|_{H^2_\mu}$ is called space $H^1(T, R)$ which is a Banach space. For $f \in H^1(T, R)$, there exists a sequence $(f_n) \subseteq C^1_\mu(T, R)$ and $h \in L^2(T, R)$ such that $\| f_n - f \|_{L^2} \rightarrow 0$, $\| f_n^\Delta - h \|^2_{L^2} \rightarrow 0$ as $n \rightarrow \infty$. The function $h$ is called the generalized derivative of $f$, denoted by $f^\Delta$. Obviously, we have $C^1_\mu(T, R) \subseteq H^1(T, R) \subseteq C^\mu(T, R) \subseteq L^2(T, R) \text{ and } C^\mu(T, R) \text{ is dense in } L^2(T, R) \text{ (see [15]).}$

Here we recall some important properties on $H^1(T, R)$ which is useful in the sequel.
Lemma 2.1 (Theorem 4.5 and Corollary 4.6 in [15]).

(1) Integration by parts in $H^1(T, R)$:
$$\int_{[s, t]} f^\Delta h (\tau) \Delta \tau + \int_{[s, t]} f^n h^\Delta (\tau) \Delta \tau = [f(\tau) h(\tau)]^t_s, \quad s, t \in T.$$  

(2) If $f, h \in H^1(T, R)$, then $fh \in H^1(T, R)$, and
$$\int_{[s, t]} (fh)^\Delta = f^\Delta h + f^n h^\Delta, \quad s, t \in T.$$  

(3) If $f \in H^1(T, R)$, then $f \in C_{id}(T, R)$ and
$$f(t) - f(s) = \int_{[s, t]} f^\Delta (\tau) \Delta \tau, \quad s, t \in T.$$  

For $f \in L^1(T, R)$, define the integral function with variable upper limit given by
$$\bar{f}(t) = \int_{[a, t]} f(\tau) \Delta \tau, \quad t \in T.$$  

Lemma 2.2 (Lemma 3.2, Lemma 3.3 and Theorem 4.7 of [15]).

(1) If $f \in C_{id}(T, R)$, then $\bar{f} \in C^1_{id}(T, R)$ and $f^\Delta (t) = \bar{f}(t), t \in T.$

(2) If $f \in L^2(T, R)$, then $\bar{f} \in H^1(T, R)$.

3. The existence of weak solution for linear dynamic equations on time scales

In order to study optimal control problems and the derive necessary conditions of optimality we have to consider the weak solution of linear dynamic equation on time scales for Cauchy problem and backward problem in particular.

First consider the following Cauchy problem
\begin{align*}
\begin{cases}
x^\Delta(t) + p(t)x^n(t) = f(t), & t > a, \\
x(a) = x_0,
\end{cases}
\end{align*}

(3.1)

where $p \in \Gamma(T, R) = \{p \in C_{id}(T, R) | 1 + \mu(t)p(t) \neq 0 \text{ for all } t \in T\}.$

A function $x \in C^1_{id}(T, R)$ is said to be a classical solution of (3.1) if $x(a) = x_0$ and for any $t \in T \setminus \{a\}$, $x$ satisfies
$$x^\Delta(t) + p(t)x^n(t) = f(t).$$

It is well known that for $f \in C_{id}(T, R)$, the classical solution $x$ of (3.1) can be by expressed
$$x(t) = e_{\otimes p}(t, a)x_0 + \int_{[a, t]} e_{\otimes p}(t, \tau)f(\tau) \Delta \tau, \quad t \in T.$$  

For $f \in L^2(T, R)$, the integral
$$\int_{[a, t]} e_{\otimes p}(t, \tau)f(\tau) \Delta \tau$$
is well defined. It is nature to define the weak solution of (3.1) given by
$$x(t) = e_{\otimes p}(t, a)x_0 + \int_{[a, t]} e_{\otimes p}(t, \tau)f(\tau) \Delta \tau, \quad t \in T.$$  

The weak solution has the following properties.

Theorem 3.1. Let $p \in \Gamma(T, R), f \in L^2(T, R)$, for every $x_0 \in R$, the weak solution $x$ of (3.1) satisfies $x \in C_{id}(T, R) \cap H^1(T, R)$ and $x^\Delta = -px^n + f$.

Proof. Since $C_{id}(T, R)$ is dense in $L^2(T, R)$, there exists a sequence $\{f_n\} \subseteq C_{id}(T, R)$ such that $f_n \longrightarrow f$ in $L^2(T, R)$ as $n \rightarrow +\infty$. Considered the following equation
\begin{align*}
\begin{cases}
x^\Delta_n(t) + p(t)x^n_n(t) = f_n(t), & t > a, \\
x_n(a) = x_0.
\end{cases}
\end{align*}

(3.2)
The Eq. (3.2) has a unique classical solution \( x_n \in C^1_{\Gamma}(T, R) \) given by

\[
x_n(t) = e_{\Gamma}(t, a)x_0 + \int_{[a, t]} e_{\Gamma}(t, s)f_n(s)\Delta s.
\]

Since \( p \in \Gamma(T, R) \), there exist \( \alpha > 0, \beta > 0 \) such that \( |p(t)| \leq \alpha, |e_{\Gamma}(t, s)| \leq \beta \) for all \( t, s \in T \). Further,

\[
\|x_n - x\|_{L^2}^2 \leq \beta^2 \int_{[a, b]} \left| \int_{[a, s]} \left( f_n(\tau) - f(\tau) \right) \Delta \tau \right|^2 \Delta t \leq \beta^2 (b-a)^3 \|f_n - f\|_{L^2}^2,
\]

\[
\|x_n - x\|_{L^2}^2 \leq \beta^2 \int_{[a, b]} \left| \int_{[a, \sigma(t)]} \left( f_n(\tau) - f(\tau) \right) \Delta \tau \right|^2 \Delta t \leq \beta^2 (b-a)^3 \|f_n - f\|_{L^2}^2,
\]

\[
\|x_n^\Delta - (-p x^\Delta + f)\|_{L^2} \leq \|p (x_n - x)\|_{L^2} + \|f_n - f\|_{L^2} \leq \left[ \alpha \beta (b-a)^2 + 1 \right] \|f_n - f\|_{L^2}.
\]

This means that \( x \in H^1(T, R) \) and \( x^\Delta = -p x^\Delta + f \). \( \square \)

Before we discuss the backward problem which is the key for adjoint equation we introduce the integral function with variable lower limit given by

\[
f(t) = \int_{[t, b]} f(\tau) \Delta \tau, \quad t \in T,
\]

for \( f \in L^2(T, R) \) and show that \( f \) has some properties similar to the integral function with variable upper limit. However, the integral function with variable lower limit on time scales cannot be turned into the integral function with variable upper limit by simple transformation \( s = b - t \).

**Lemma 3.1.**

1. If \( f \in C_{\Gamma}(T, R) \), then \( f^\Delta \in C^1_{\Gamma}(T, R) \) and

\[
f^\Delta(t) = -f(t) \quad \text{for all } t \in T^k.
\]

2. If \( f \in L^2(T, R) \), then \( f \in H^1(T, R) \).

**Proof.**

1. If \( t \) is right-scattered i.e. \( t < \sigma(t) \), by **Definition 2.1** we have

\[
\frac{f(\sigma(t)) - f(t)}{\sigma(t) - t} = \frac{1}{\sigma(t) - t} \left[ \int_{[t, \sigma(t)]} f(\tau) d\tau - f(t) \right] = -\frac{1}{\sigma(t) - t} \int_t^{\sigma(t)} f(\tau) d\tau = -f(t).
\]

If \( t \) is right-dense i.e. \( t = \sigma(t) \) in addition \( s \neq t \), then

\[
\frac{f(\sigma(t)) - f(s)}{\sigma(t) - s} = \frac{1}{t - s} \left[ \int_s^b f(\tau) d\tau - f(t) \right] = -\frac{1}{t - s} \int_s^t f(\tau) d\tau.
\]

By the stand Lebesgue integral theory, \( \int_s^t f(\tau) d\tau \) is a continuous function on the interval \([a, b]\), we have

\[
\lim_{s \to t} \frac{1}{t - s} \int_s^t f(\tau) d\tau = f(t).
\]

Therefore \( f^\Delta \in C^1_{\Gamma}(T, R) \) and

\[
f^\Delta(t) = -f(t) \quad \text{for any } t \in T^k.
\]

2. Since the space \( C_{\Gamma}(T, R) \) is dense in \( L^2(T, R) \), for any \( f \in L^2(T, R) \), there is a sequence \( \{f_n\} \subseteq C_{\Gamma}(T, R) \) such that \( f_n \to f \) in \( L^2(T, R) \) as \( n \to +\infty \). Let

\[
f_n(t) = \int_{[t, b]} f_n(\tau) \Delta \tau, \quad t \in T,
\]

we have \( f_n \in C^1_{\Gamma}(T, R), f_n^\Delta(t) = f_n(t) \)

\[
\|f_n - f\|_{L^2}^2 = \int_{[a, b]} \left| \int_{[t, b]} (f_n(\tau) - f(\tau)) \Delta \tau \right|^2 \Delta t \leq \int_{[a, b]} (b-t)^2 \|f_n - f\|_{L^2}^2 \Delta t \leq (b-a)^3 \|f_n - f\|_{L^2}^2.
\]

Thus \( f \in H^1(T, R) \). \( \square \)
Now, we consider the following backward problem

\[
\begin{align*}
\varphi^\Delta(t) &= p(t)\varphi(t) + h(t), \\
\varphi(b) &= \varphi_0, \\
\end{align*}
\]  
(3.3)

where \( p \in \Gamma(T, R) \).

**Definition 3.1.** A function \( \varphi \in C^1_{ad}(T, R) \) is said to be a classical solution of the Eq. (3.3), if \( \varphi \) satisfies the following conditions:

1. \( \varphi(b) = \varphi_0 \),
2. \( \varphi^\Delta(t) = p(t)\varphi(t) + h(t) \) for all \( t \in T^k \).

**Theorem 3.2.** If \( h \in C_{ad}(T, R) \), then the Eq. (3.3) has a unique classical solution \( \varphi \in C^1_{ad}(T, R) \) given by

\[
\varphi(t) = e_{\varphi p}(b, t)\varphi_0 - \int_{[t,b)} e_{\varphi p}(s, t)h(s)\Delta s.
\]  
(3.4)

**Proof.** First, using the properties of exponential function, we have

\[
(e_{\varphi p}(b, t))^\Delta = p(t)e_{\varphi p}(b, t).
\]

Define

\[
\gamma(t) = \int_{[t,b)} e_{\varphi p}(s, t)h(s)\Delta s.
\]

By the properties of the integral function with variable lower limit (see Lemma 3.1) and elementary computation one can verify

\[
\gamma^\Delta(t) = p(t)\gamma(t) - h(t).
\]

Thus

\[
\varphi^\Delta(t) = p(t)\varphi(t) + h(t).
\]

This means that the function \( \varphi \in C^1_{ad}(T, R) \) and it just is the classical of the Eq. (3.3).

On the other hand, if the function \( \psi \) is a classical solution of the Eq. (3.3), then

\[
[e_{p}(b, \cdot)\psi(\cdot)]^\Delta = -p(\cdot)e_{p}(b, \sigma(\cdot))\psi(\cdot) + e_{p}(b, \sigma(\cdot))\psi^\Delta(\cdot).
\]

We integrate both sides form \( t \) to \( b \) to conclude

\[
\psi(t) = e_{\varphi p}(b, t)\varphi_0 - \int_{[t,b)} e_{\varphi p}(\sigma(t), t)h(t)\Delta t.
\]

This implies the solution is unique. \( \square \)

By discussion similar to the Cauchy problem, for \( p \in \Gamma(T, R), g \in L^2(T, R) \), we can introduce the weak solution of (3.3) given by

\[
\varphi(t) = e_{\varphi p}(b, t)\varphi_0 - \int_{[t,b)} e_{\varphi p}(s, t)h(s)\Delta s, \quad t \in T,
\]

and verify the following theorem.

**Theorem 3.3.** Suppose that \( p \in \Gamma(T, R), h \in L^2(T, R), \varphi_0 \in R \), then the Eq. (3.1) has a unique weak solution \( \varphi \) and \( \varphi \in C_{ad}(T, R) \cap H^1(T, R), \varphi^\Delta = p\varphi + h \).

4. **Necessary conditions of optimality**

First, we formulate optimal control problems on time scales. Assume that the control space is given by the Hilbert space \( L^2(T, R) \). The class of admissible control is given by a nonempty closed and convex subset \( U_{ad} \) of \( L^2(T, R) \). Consider the controlled system

\[
\begin{align*}
x^\Delta(t) + p(t)x(t) &= f(t) + u(t), \\
x(a) &= x_0, \\
u &\in U_{ad},
\end{align*}
\]  
(4.1)

where \( p \in \Gamma(T, R) \). By Theorem 3.1, one can easily establish the following theorem.
Theorem 4.1. Let \( p \in \Gamma(T, R) \), \( f \in L^2(T, R) \), for every \( u \in U_{ad} \), the controlled system (4.1) has a unique weak solution \( x \) given by

\[
x(t) = e_{\mathcal{L}^p}(t, a)x_0 + \int_{[a, t]} e_{\mathcal{L}^p}(t, \tau)[f(\tau) + u(\tau)]\Delta \tau, \quad t \in T.
\]

The Lagrange problem (P): Find \( \bar{u} \in U_{ad} \) such that

\[
J(\bar{u}) \leq J(u) \quad \text{for all} \quad u \in U_{ad},
\]

where \( J \) is the cost functional given by

\[
J(u) = \int_{[a, b]} |x'(t) - x_d(t)|^2 \Delta t + \int_{[a, b]} |u(t)|^2 \Delta t
\]

and \( x_d \in L^2(T, R) \) is a desired value, \( x \in C_{ad}(T, R) \cap H^1(T, R) \) denotes the weak solution of (4.1) corresponding to the control \( u \in U_{ad} \).

Now, we present necessary conditions for optimality.

Theorem 4.2. Suppose that \( f \in L^2(T, R) \), \( p \in \Gamma(T, R) \), \( x_d \in L^2(T, R) \), then, in order that the pair \( (\bar{x}, \bar{u}) \) be optimal, it is necessary that there exists a function \( \varphi \in H^1(T, R) \) such that the following equations and inequality hold:

\[
\begin{aligned}
\dot{\varphi}^a(t) &= p(t)\varphi(t) - [\bar{x}'(t) - x_d(t)], \\
\varphi(b) &= 0.
\end{aligned}
\]

\[
\int_{[a, b]} (\bar{u}(t) + \varphi(t)) (u(t) - \bar{u}(t)) \Delta t \geq 0.
\]

**Proof.** Since \((\bar{x}, \bar{u}) \in H^1(T, R) \times U_{ad}\) is an optimal pair, it must satisfies the Eq. (4.2).

Since \(U_{ad}\) is convex, it is clear that \( u_\epsilon = \bar{u} + \epsilon (u - \bar{u}) \in U_{ad} \) for \( \epsilon \in [0, 1] \), for \( u \in U_{ad} \). Let \( x_\epsilon \) be the weak solution of (4.1) corresponding to the control \( u_\epsilon \), then \( x_\epsilon \) can be expressed by

\[
x_\epsilon(t) = e_{\mathcal{L}^p}(t, a)x_0 + \int_{[a, t]} e_{\mathcal{L}^p}(t, s) [f(s) + u_\epsilon(s)] \Delta s.
\]

Considering

\[
x_\epsilon(t) - \bar{x}(t) = \int_{[a, t]} e_{\mathcal{L}^p}(t, s) [u_\epsilon(s) - \bar{u}(s)] \Delta s = \epsilon \int_{[a, t]} e_{\mathcal{L}^p}(t, s) [u(s) - \bar{u}(s)] \Delta s.
\]

and set

\[
y = \lim_{\epsilon \to 0} \frac{x_\epsilon - \bar{x}}{\epsilon},
\]

we have

\[
y(t) = \int_{[a, t]} e_{\mathcal{L}^p}(t, s) [u(s) - \bar{u}(s)] \Delta s.
\]

By virtue of **Theorem 3.1**, it is easy to see that the function \( y \) is just the weak solution of the following variational equation

\[
\begin{aligned}
\dot{y}^a(t) + p(t)y^a(t) &= u(t) - \bar{u}(t), \\
y(a) &= 0.
\end{aligned}
\]

In fact, \( y \) is the Gateaux derivative of weak solution \( x \) at \( \bar{u} \) in the direction \( u - \bar{u} \).

Consider the adjoint equation

\[
\begin{aligned}
\dot{\varphi}^a(t) &= p(t)\varphi(t) - [\bar{x}'(t) - x_d(t)], \\
\varphi(b) &= 0.
\end{aligned}
\]

By virtue of **Theorem 3.3**, the Eq. (4.8) has a unique weak solution \( \varphi \in C_{rd}(T, R) \cap H^1(T, R) \) given by

\[
\varphi(t) = \int_{[t, b]} e_{\mathcal{L}^p}(s, t) [\bar{x}'(s) - x_d(s)] \Delta s.
\]
By the integration by parts in $H^1(T, R)$ (see Lemma 2.1), we have
\[
\int_{[a,b]} y^\sigma (t) \{ \bar{x}^\sigma (t) - x_d(t) \} \Delta t = \int_{[a,b]} y^\sigma (t) \{ -\varphi^\Delta (t) + p(t)\varphi (t) \} \Delta t
\]
\[
= \int_{[a,b]} \varphi (t) \{ y^\sigma (t) + p(t)y^\sigma (t) \} \Delta t
\]
\[
= \int_{[a,b]} \varphi (t) \{ u(t) - \bar{u}(t) \} \Delta t. \tag{4.9}
\]
\(\bar{u}\) is the optimal control, hence
\[
J(u_0) - J(\bar{u}) \geq 0. \tag{4.10}
\]
Using (4.5), (4.6) and (4.9), we obtain
\[
J(u_0) - J(\bar{u}) = \int_{[a,b]} [u^\sigma (t) + \bar{u}(t)] [u^\sigma (t) - \bar{u}(t)] \Delta t + \int_{[a,b]} [x^\sigma _x (t) - \bar{x}^\sigma (t)] [\bar{x}^\sigma _x (t) + \bar{x}^\sigma (t) - 2x_d(t)] \Delta t
\]
\[
= 2\epsilon \int_{[a,b]} \bar{u}(t) [u(t) - \bar{u}(t)] \Delta t + \epsilon^2 \int_{[a,b]} |u(t) - \bar{u}(t)|^2 \Delta t
\]
\[
+ \epsilon \int_{[a,b]} y^\sigma (t) [x^\sigma _{xx} (t) + \bar{x}^\sigma (t) - 2x_d(t)] \Delta t
\]
\[
= 2\epsilon \int_{[a,b]} \bar{u}(t) [u(t) - \bar{u}(t)] \Delta t + 2\epsilon \int_{[a,b]} y^\sigma (t) [\bar{x}^\sigma (t) - x_d(t)] \Delta t
\]
\[
+ \epsilon^2 \int_{[a,b]} |y^\sigma (t)|^2 \Delta t + \epsilon^2 \int_{[a,b]} |u(t) - \bar{u}(t)|^2 \Delta t
\]
\[
= 2\epsilon \int_{[a,b]} [\bar{u}(t) + \varphi (t)] [u(t) - \bar{u}(t)] \Delta t + \epsilon^2 \int_{[a,b]} [|y^\sigma (t)|^2 + |u(t) - \bar{u}(t)|^2] \Delta t. \tag{4.11}
\]
Substituting (4.11) into (4.10) we have the following inequality
\[
\int_{[a,b]} [\bar{u}(t) + \varphi (t)] [u(t) - \bar{u}(t)] \Delta t \geq 0. \tag{4.12}
\]
This completes the proof of all the necessary conditions as stated. \(\square\)

As contrast, we can consider the following optimal control problem (P1): Find $\bar{u} \in U_{ad}$ such that
\[
J(\bar{u}) \leq J(u) \quad \text{for all } u \in U_{ad}.
\]
Subject to
\[
\begin{cases}
  x^\Delta (t) = p(t)x(t) + f(t) + u(t), \\
  x(a) = x_0, \\
  u \in U_{ad},
\end{cases} \tag{4.13}
\]
where $p \in \Gamma'(T, R), f$ is the cost functional given by
\[
J(u) = \int_{[a,b]} |x(t) - x_d(t)|^2 \Delta t + \int_{[a,b]} |u(t)|^2 \Delta t
\]
and $x_d \in L^2(T, R)$ is a desired value.

By procedures similar to Theorem 4.2, one can derive the necessary conditions of optimality for the problem (P1).

**Theorem 4.3.** Suppose that $f \in L^2(T, R), p \in \Gamma'(T, R), x_d \in L^2(T, R)$, then, in order that the pair $((\bar{x}, \bar{u})$ be optimal, it is necessary that there exists a function $\varphi \in H^1(T, R)$ such that the following equations and inequality hold:
\[
\begin{cases}
  \bar{x}^\Delta (t) = p(t)\bar{x}(t) + f(t) + \bar{u}(t), \\
  \bar{x}(a) = x_0, \\
  \varphi^\Delta (t) + p(t)\varphi^\sigma (t) = -[\bar{x}(t) - x_d(t)], \\
  \varphi(b) = 0, \\
  \int_{[a,b]} [\bar{u}(t) + \varphi^\sigma (t)] [u(t) - \bar{u}(t)] \Delta t \geq 0.
\end{cases} \tag{4.14-16}
\]
Now, we can reveal relationship and difference between the problem \( (\text{P}) \) and \( (\text{P1}) \) in the following table:

<table>
<thead>
<tr>
<th>Item</th>
<th>Problem ((\text{P}))</th>
<th>Problem ((\text{P1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled system</td>
<td>( x^{(4)}(t) + p(t)x''(t) = f(t) + u(t), )</td>
<td>( x^{(4)}(t) = p(t)x(t) + f(t) + u(t), )</td>
</tr>
<tr>
<td>Weak solution of controlled system</td>
<td>( x(t) = e_{\phi}(t, a)x_0 + \int_{a}^{t} e_{\phi}(t, s)[f(s) + u(s)]ds )</td>
<td>( x(t) = e_{\phi}(t, a)x_0 + \int_{a}^{t} e_{\phi}(t, s)f(s) + u(s)\Delta s )</td>
</tr>
<tr>
<td>Cost functional</td>
<td>( f(u(a), b) = 0 )</td>
<td>( f(u(a), b) = 0 )</td>
</tr>
<tr>
<td>Adjoint equation</td>
<td>( \psi^{(2)}(t) + p(t)\psi(t) = -[\ddot{x}(t) - \chi(t)], )</td>
<td>( \psi^{(2)}(t) + p(t)\psi(t) = -[\ddot{x}(t) - \chi(t)], )</td>
</tr>
<tr>
<td>Weak solution of adjoint equation</td>
<td>( \psi(3) = 0 )</td>
<td>( \psi(3) = 0 )</td>
</tr>
<tr>
<td>Optimal inequality</td>
<td>( \int_{a}^{b} [\ddot{u}(t) + \psi(t)]</td>
<td>u(t) - \ddot{u}(t)</td>
</tr>
</tbody>
</table>

5. Examples

\( T = R \) and \( T = Z \) are the most typical examples of time scales corresponding to differential and difference equations, respectively.

**Example 1.** Let \( T = [0, 3] \subset R \) be real interval, consider the following dynamic equation

\[
\begin{align*}
\ddot{x}(t) + tx(t) &= u(t), \quad t > 0, \\
x(0) &= x_0, \\
u &\in L^2(T, R)
\end{align*}
\] (5.1)

with the cost functional

\[
J(u) = \int_{0}^{3} (|x(t)|^2 + |u(t)|^2) \, dt.
\]

For **Example 1**, using the necessary conditions of optimality given by **Theorem 4.2**, one can show the following theorem.

**Theorem 5.1.** In order to that the pair \((\ddot{x}, \ddot{u})\) be optimal, it is necessary that there exists a function \( \psi \in H^1(T, R) \) such that the equations and inequality hold:

\[
\begin{align*}
\ddot{x}(t) + t\dot{x}(t) &= \ddot{u}(t), \quad t \in (0, 3] \\
x(0) &= x_0, \\
\dot{\psi}(t) &= \dot{\psi}(t) - \dddot{x}(t), \quad t \in [0, 3) \\
\psi(3) &= 0, \\
\int_{0}^{3} [\dddot{u}(t) + \dot{\psi}(t)]|u(t) - \dddot{u}(t)| \, dt &\geq 0.
\end{align*}
\] (5.2)

In this case the dynamic equation on \( T \) is just ordinary differential equation. The result coincides with the classical conclusion for optimal control problems of systems governed by the ordinary differential equation (see Section 6.3.2 of [16]).

**Example 2.** We consider the quantum time scale \( T = \{t = q^k | q > 1, 0 \leq k \leq N_0\} \). Obviously, we have \( \sigma(t) = qt, \mu(t) = (q - 1)t \). Let

\[
U_{ad} = \left\{ u : T \rightarrow \mathbb{R} | \sum_{k=0}^{N_0} \mu(k) \| u(q^k) \| < +\infty \right\}, \quad \| u \| = \sum_{k=0}^{N_0} \mu(k) |u(q^k)|,
\]

one can verify that \( U_{ad} \) is a nonempty closed convex set. Study the following dynamic equation

\[
\begin{align*}
x(q^{k+1}) &= \frac{1}{1 + \mu(q^k)} p(q^k)x(q^k) + \frac{\mu(q^k)}{1 + \mu(q^k)} p(q^k)u(q^k), \quad 0 \leq k \leq N_0 - 1, \\
x(1) &= x_0, \\
u &\in U_{ad},
\end{align*}
\] (5.5)

with the cost functional

\[
J(u) = \sum_{k=0}^{N_0-1} |x(q^{k+1})|^2 + \sum_{k=0}^{N_0} |u(q^k)|^2.
\]

For this example, our results can be used to (5.5), that is, we have the following theorem.
Theorem 5.2. In order that the pair \( \{ \bar{x}, \bar{u} \} \) is optimal, it is necessary that there exists a sequence \( \{ \psi (q^k) \} \) such that

\[
\begin{align*}
\bar{x}(q^{k+1}) &= \frac{1}{1 + \mu (q^k) p(q^k)} \bar{x}(q^k) + \frac{\mu (q^k)}{1 + \mu (q^k) p(q^k)} \bar{u}(q^k), \quad 0 \leq k \leq N_0 - 1, \\
\bar{x}(1) &= x_0.
\end{align*}
\] (5.6)

\[
\begin{align*}
\psi (q^k) &= \frac{1}{1 + \mu (q^k) p(q^k)} [\psi (q^{k+1}) + \mu (q^k) \bar{x}(q^{k+1})], \quad 0 \leq k \leq N_0 - 1, \\
\psi (q^{N_0}) &= 0.
\end{align*}
\] (5.7)

\[
\sum_{k=0}^{N_0} [\bar{u}(q^k) + \psi (q^k)] [u(q^k) - \bar{u}(q^k)] \geq 0.
\] (5.8)

Example 3. Consider mathematical programming problem

\[
\min \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij}(x_{ij})
\]

s.t.

\[
\begin{align*}
\sum_{j=1}^{m} x_{ij} &\leq a_i, \quad i = 1, 2, \ldots, n, \\
\sum_{i=1}^{n} x_{ij} &\geq b_j, \quad j = 1, 2, \ldots, m, \\
x_{ij} &\geq 0, \quad i = 1, 2, \ldots, n; j = 1, 2, \ldots, m,
\end{align*}
\] (5.9)

where \( \sum_{i=1}^{n} a_i \) and \( \sum_{j=1}^{m} b_j \) are positive. Firstly, the mathematical programming problem (5.9) be rewritten as optimal control problem. Let \( T = \{0, 1, 2, \ldots, m\} \).

\[
x(0) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \quad x(1) = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}, \quad x(k) = \begin{pmatrix} \sum_{j=1}^{k} x_{ij} \\ \vdots \\ \sum_{j=1}^{k} x_{nj} \end{pmatrix},
\]

\[
u(k) = \begin{pmatrix} u_1(k) \\ \vdots \\ u_n(k) \end{pmatrix}, \quad x_{1,k+1} = \begin{pmatrix} x_{1,1} \\ \vdots \\ x_{1,n} \\ x_{1,k+1} \end{pmatrix}, \quad C_{k+1}(u(k)) = \begin{pmatrix} c_{1,1,k+1} (u_1(k)) \\ \vdots \\ c_{n,1,k+1} (u_n(k)) \end{pmatrix},
\]

\[
U_{ad} = \left\{ u : T \rightarrow \mathbb{R}^n | 0 \leq u(k) \leq \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \sum_{j=0}^{k-1} u(j), (1 \cdots 1) u(k) \geq b_{k+1} \right\}.
\]

Obviously, \( U_{ad} \neq \emptyset \) is nonempty closed convex and

\[
\begin{align*}
x(k+1) &= x(k) + u(k), \quad k = 0, 1, 2, \ldots, m - 1 \\
x(0) &= 0, \\
u &\in U_{ad}.
\end{align*}
\] (5.10)

Define cost functional

\[
J(u) = \sum_{k=0}^{m-1} C_{k+1} (u(k)).
\]

Optimal control problem (P): Find \( \bar{u} \in U_{ad} \) such that

\[
J (\bar{u}) \leq J(u) \quad \text{for all } u \in U_{ad}.
\]

Theorem 5.3. Suppose that \( C_k : \mathbb{R}^n \rightarrow \mathbb{R} \) is convex \( k = 1, 2, \ldots, m \). In order that the pair \( \{ \bar{x}, \bar{u} \} \) is optimal, it is necessary that the following equality and inequality hold:
\begin{equation}
\begin{aligned}
\tilde{x}(k+1) &= \tilde{x}(k) + \bar{u}(k), \quad 0 \leq k \leq m - 1, \\
\tilde{x}(0) &= 0.
\end{aligned}
\tag{5.11}
\end{equation}

\begin{equation}
\sum_{k=0}^{m-1} C_{k+1}^i(\bar{u}(k)) [u(k) - \bar{u}(k)] \geq 0.
\tag{5.12}
\end{equation}

Here, it is easy to see for (5.12) that \( C_{k+1}^i(\bar{u}(k)) \) is just the value of Lagrange multiplier \( k = 0, 1, \ldots, m - 1 \). Specially, consider the following nonlinear mathematical programming problem

\[\min \ x_1^2 + x_2^2 + x_3^2 \]

\[\text{s.t.} \quad x_1 + x_2 + x_3 = 100. \tag{5.13}\]

Let \( T = \{0, 1, 2\} \), \( \bar{x}(0) = 0, x(1) = x_1, x(2) = x_1 + x_2, x(3) = x_1 + x_2 + x_3, u(0) = x_1, u(1) = x_2u(2) = x_3 \). By the value of Theorem 5.3, \( (\bar{x}, \bar{u}) \) is optimal, it is necessary that the following equality and inequality hold:

\[\begin{aligned}
\tilde{x}(k+1) &= \tilde{x}(k) + \bar{u}(k), \\
\bar{x}(0) &= 0, \\
2\bar{u}(0) (u(0) - \bar{u}(0)) + 2\bar{u}(1) (u(1) - \bar{u}(1)) + 2\bar{u}(2) (u(2) - \bar{u}(2)) &\geq 0.
\end{aligned}\]

Furthermore, we immediately obtain \( \bar{u}(0) = \bar{u}(1) = \bar{u}(2) = \frac{100}{3} \).

Acknowledgments

The author expresses his sincere thanks to the referees of this paper. Many suggestions due to them were quite valuable for revising the manuscripts in the present style.

This work is supported by the National Natural Science Foundation of China under grant no. 10661004 and the National Natural Science and Technology Committee of Guizhou province under grant no. 20092069.

References