Automatica 47 (2011) 748–753

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

# Initial condition of costate in linear optimal control using convex analysis☆

Bin Liu *

China Mobile Group, Tianjin Co., Ltd., Tianjin, China

ARTICLE INFO

*Article history:*
Available online 11 March 2011

*Keywords:*
Pontryagin maximum principle
Optimal control
Convex analysis
Duality

ABSTRACT

The Pontryagin Maximum Principle is one of the most important results in optimal control, and provides necessary conditions for optimality in the form of a mixed initial/terminal boundary condition on a pair of differential equations for the system state and its conjugate costate. Unfortunately, this mixed boundary value problem is usually difficult to solve, since the Pontryagin Maximum Principle does not give any information on the initial value of the costate. In this paper, we explore an optimal control problem with linear and convex structure and derive the associated dual optimization problem using convex duality, which is often much easier to solve than the original optimal control problem. We present that the solution to the dual optimization problem supplements the necessary conditions of the Pontryagin Maximum Principle, and elaborate the procedure of constructing the optimal control and its corresponding state trajectory in terms of the solution to the dual problem.



## 1. Introduction

The calculus of variations is a branch of mathematics which generalizes ordinary calculus, and seeks to find the path, curve, surface, etc., for which a given function has a stationary value. The origin of the calculus of variations can be traced back to the 17th century works of Bernoulli and Newton. It was developed further in the 18th century by Euler and Lagrange and in the 19th century by Jacobi, Hamilton and Weierstrass. In the early 20th century, Bolza and Bliss gave the calculus of variations its present rigorous mathematical structure on the basis of Weierstrass's work. Optimal control is an extension of the calculus of variations, with the goal of finding the input control function which minimizes a given cost functional with differential equation constraints (Bryson, 1996). The maximum principle formulated in the 1950s by the Russian mathematicians Pontryagin, Boltyanskii, Gamkrelidze and Mishchenko is by far the most important result in optimal control theory, and marked the emergence of optimal control as a distinct field of research (Pontryagin, Boltyanskii, Gamkrelidze, & Mishchenko, 1962). The Pontryagin Maximum Principle consists of a system of state differential equations with initial condition and a corresponding system of costate differential equations with terminal condition, and not only includes every single known result in the calculus of variations as special cases, but also allows us to tackle optimization problems beyond the reach of all other methods. The lack of initial condition of the costate makes the two point boundary value problem rather difficult to handle in general.

Several numerical techniques without requirement of the initial costate have been proposed in the literature. In Teo, Goh, and Wong (1991), the original optimal control problem is approximated by a sequence of optimal parameter selection problems through control parametrization, where the control functions are approximated by piecewise constant functions. Consequently, the optimal control problem can readily be treated as nonlinear programming problems if the gradients of the cost functional with respect to the decision parameters are obtained. However, computation of the required gradients involves the integration of the state differential equations forward in time, followed by the integration of the corresponding costate differential equations backward in time. The state and costate systems are solved in opposite directions using an adaptive integration scheme, hence it is impossible to ensure that the state and costate knot sets coincide. Furthermore, since solving the costate system is dependent on the solution of the state system, an appropriate interpolation method needs to be invoked, which compromises the accuracy of the resulting gradients. A novel alternative scheme is introduced by virtue of a new auxiliary system of differential equations in replacement of the costate system Loxton, Teo, and Rehbock (2008), where the auxiliary system can be solved simultaneously with the state system.

The initial value of the costate can be furnished in some special optimal control problems. For the linear quadratic regulator with a finite time horizon, a well known procedure is available for retrieving the initial condition based on linearity of the ordinary differential equations in the maximum principle (Sontag, 1998). Successively solving two first order quasilinear partial differential

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Kok Lay Teo under the direction of Editor Ian R. Petersen.

* Tel.: +86 22 84407115.
*E-mail address:* liubinh@yahoo.com.

0005-1098/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2011.01.033

equations yields the initial costate for the nonlinear quadratic problem (Costanza, 2007; Costanza & Rivadeneira, 2008).

In this paper, we only consider a simple representative optimal control problem, known as a convex Mayer problem with linear dynamics, and propose an algorithm to provide the initial condition of the costate and to give the optimal control and its corresponding state trajectory in an analytical way without solving the maximum principle numerically. The paper is organized as follows. In Section 2, the maximum principle is applied to the Mayer problem, where the deficiency of the maximum principle is illustrated. Next we briefly review convex analysis, the main tool used to overcome the deficiency, in Section 3. In Section 4, the Mayer problem is formulated as the Bolza problem in the calculus of variations. We then explore the dual optimization problem of the resulting Bolza problem by taking advantage of convex analysis. The dual problem can be transformed into a minimization problem over the finite dimensional Euclidean space without any constraints, whose solution is the desired initial value of the costate and the key to finding optimal control and its corresponding state trajectory. Finally we give some conclusions in the last section.

## 2. Problem statement

Consider a process evolving over the fixed time horizon $[0, T]$, where $T$ is a given terminal time, and satisfying the dynamical system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{1}$$

with the initial condition

$$x(0) = x_0. \tag{2}$$

Here, $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$ are the state and input control of the system at time $t$, respectively. $A(t)$ and $B(t)$ are real matrices of dimensions of $n \times n$ and $n \times m$, respectively. Let $U$ be a convex and compact subset of $\mathbb{R}^m$. Any Borel measurable function $u : [0, T] \longrightarrow U$ such that $u(t) \in U$ almost everywhere on $[0, T]$ is said to be an admissible control. Let $\mathcal{U}$ be the class of all such admissible controls. For a given $u \in \mathcal{U}$, the dynamical system (1) has a unique absolutely continuous solution $x(\cdot|u)$ that satisfies the initial condition (2). With this groundwork, we are now able to introduce the linear/convex Mayer problem.

**Problem** (P) Given the dynamical system (1) and the initial condition (2), find a $u \in \mathcal{U}$ such that the cost functional

$$g(x(T|u))$$

is minimized over $\mathcal{U}$, where $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a convex and continuously differentiable function.

The problem can be restated as finding a $u^* \in \mathcal{U}$ such that

$$g(x(T|u^*)) \le g(x(T|u))$$

for all $u \in \mathcal{U}$. If $u^* \in \mathcal{U}$ is an optimal control, and $x^*(t)$ and $\lambda^*(t)$ are the corresponding state and costate, then from the maximum principle,

$$\dot{\lambda}^*(t) = -A'(t)\lambda^*(t),$$

with the terminal condition

$$\lambda^*(T) = -\nabla g(x^*(T)),$$

where $A'(t)$ is the transpose of $A(t)$, and $\lambda^* : [0, T] \longrightarrow \mathbb{R}^n$ is an absolutely continuous function. Moreover,

$$\begin{aligned}\langle \lambda^*(t), A(t)x^*(t) + B(t)u^*(t)\rangle &= \max_{u\in U}\langle \lambda^*(t), A(t)x^*(t) + B(t)u\rangle \\ &= \langle \lambda^*(t), A(t)x^*(t)\rangle + \max_{u\in U}\langle \lambda^*(t), B(t)u\rangle,\end{aligned}$$

where "$\langle \cdot, \cdot\rangle$" represents the inner product. Thus,

$$\begin{aligned}\langle \lambda^*(t), B(t)u^*(t)\rangle &= \max_{u\in U}\langle \lambda^*(t), B(t)u\rangle \\ &= \max_{u\in U}\langle B'(t)\lambda^*(t), u\rangle,\end{aligned}$$

i.e., $u^*(t)$ is a point in $U$ maximizing the mapping

$$u \longrightarrow \langle B'(t)\lambda^*(t), u\rangle : U \longrightarrow \mathbb{R}. \tag{3}$$

**Assumption 1.** For simplicity, we always suppose that for each $\xi \in \mathbb{R}^m$, the mapping $u \longrightarrow \langle \xi, u\rangle : U \longrightarrow \mathbb{R}$ is maximized at a single point $\gamma(\xi)$, that is

$$\arg\max_{u\in U}\langle \xi, u\rangle = \gamma(\xi), \quad \forall \xi \in \mathbb{R}^m.$$

Then from (3), the optimal control $u^*(t)$ is necessarily given by

$$u^*(t) = \gamma(B'(t)\lambda^*(t)).$$

Therefore we summarize the two point boundary value problem

$$\begin{cases}\text{(a) } \dot{x}^*(t) = A(t)x^*(t) + B(t)u^*(t), \qquad x^*(0) = x_0,\\ \text{(b) } u^*(t) = \gamma(B'(t)\lambda^*(t)),\\ \text{(c) } \dot{\lambda}^*(t) = -A'(t)\lambda^*(t),\\ \text{(d) } \lambda^*(T) = -\nabla g(x^*(T)).\end{cases} \tag{4}$$

It is well known that the necessary optimality conditions (4) for the Mayer Problem (P) are also sufficient (Azhmyakov & Raisch, 2008). Unfortunately, (4) is a difficult system of equations to handle, mainly because it involves an initial condition for the state $x^*(t)$ but only a terminal condition for the costate $\lambda^*(t)$. In particular, the maximum principle says nothing about what initial value of the costate should be used with (4)(c) in order that the remaining conditions of (4) be satisfied. A computational method is developed in Teo et al. (1991), where the Mayer Problem (P) is approximated by a sequence of parameter selection problems using control parametrization such that the solution to each of these approximate problems is a suboptimal solution to the Mayer Problem (P). Furthermore, the sequence of controls generated by the method converges to the true optimal control in the weak* topology of $L_\infty([0, T], \mathbb{R}^m)$. Nevertheless, the system (4) is numerically unstable when the costate system is solved forward in time unless an accurate initial condition is available. To overcome the deficiency, in what follows we derive an algorithm to furnish the initial costate by exploiting the special structure of our problem and convex duality.

## 3. Preliminaries from convex analysis

In this section we briefly discuss the generalized Bolza problem in the calculus of variations, which is characterized by non-smooth but convex data (Clarke, 1976). Then its dual problem and several important results are presented, as have been thoroughly discussed in Rockafellar (1970a). The prerequisites turn out to yield an approach to solving the Mayer Problem (P) which offers the initial costate noted in the previous section with respect to the maximum principle. Let $l$ and $L(\cdot, \cdot, t)$ be convex and lower semicontinuous functions on $\mathbb{R}^n \times \mathbb{R}^n$ with values in $(-\infty, +\infty]$, not identically $+\infty$. Furthermore, $L$ is $\mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n) \times \mathcal{L}[0, T]$-measurable. Let $A_n^1$ represent the set of absolutely continuous functions from $[0, T] \longrightarrow \mathbb{R}^n$. The Bolza problem is stated as follows.

**Problem** ($\mathrm{P}_1$) Find an $x \in A_n^1$ such that the cost functional

$$\Phi(x) = l(x(0), x(T)) + \int_0^T L(x(t), \dot{x}(t), t)\mathrm{d}t$$

is minimized over $A_n^1$.