J Glob Optim (2012) 54:221–233
DOI 10.1007/s10898-009-9474-3

Applying the canonical dual theory in optimal control problems

Jinghao Zhu · Dan Wu · David Gao

Received: 11 September 2009 / Accepted: 21 September 2009 / Published online: 7 October 2009

Abstract This paper presents some applications of the canonical dual theory in optimal control problems. The analytic solutions of several nonlinear and nonconvex problems are investigated by global optimizations. It turns out that the backward differential flow defined by the KKT equation may reach the globally optimal solution. The analytic solution to an optimal control problem is obtained via the expression of the co-state. Some examples are illustrated.

Keywords Optimal control · Canonical dual method · Global optimization

1 Introduction

It is well known that there is a close relationship between the theory of optimization and the technique of optimal control [4,11,14]. This paper is devoted to the study of optimal control problems by the canonical dual theory which has been widely used in the research of global optimizations recently [5,7–9]. As the basic model for our study, we consider the following optimal control problem (primal problem $(\mathcal{P})$ in short):

$$(\mathcal{P}) \quad \min \int_0^T [F(x) + P(u)]dt, \tag{1.1}$$

This research was partly supported by the National Science Foundation of China under grants No.10671145.

J. Zhu (✉)
Department of Applied Mathematics, Tongji University, Shanghai, China
e-mail: jinghaok@online.sh.cn

D. Wu
Department of Mathematics, Tongji University, Shanghai, China

D. Gao
Department of Mathematics, Virginia Tech, Blacksburg, VA, USA

$$\text{s.t. } \dot{x} = A(t)x + B(t)u, \quad x(0) = x_0 \in R^n, \quad t \in [0, T], \quad \|u\| \leq 1. \tag{1.2}$$

where $F(.)$ is continuous on R^n, $P(.)$ is twice continuously differentiable in R^m. An admissible control, taking values on the unit ball $D = \{u \in R^m \mid u^T u \leq 1\}$, is integrable or piecewise continuous on $[0, T]$. In (1.2) we assume that $A(t)$, $B(t)$ are continuous matrix functions in $C([0, T], R^{n\times n})$ and $C([0, T], R^{n\times m})$, respectively. This problem often comes up as a main objective in general optimal control theory [16].

By the classical control theory [16,18], we have the following Hamilton–Jacobi–Bellman function

$$H(t, x, u, \lambda) = \lambda^T (A(t)x + B(t)u) + F(x) + P(u). \tag{1.3}$$

The state and co-state systems are as follows

$$\dot{x} = H_\lambda(t, x, u, \lambda) = A(t)x + B(t)u, \quad x(0) = x_0; \tag{1.4}$$
$$\dot{\lambda} = -H_x(t, x, u, \lambda) = -A^T\lambda - \nabla F(x), \quad \lambda(T) = \vec{0}. \tag{1.5}$$

In general, it is difficult to obtain an analytic form of the optimal feedback control for the problem (1.1)–(1.2). It is well known that, in the case of unconstraint, when $P(u)$ is a positive definite quadratic form and $F(x)$ is a positive semi-definite quadratic form, we have a perfect optimal feedback control for the problem. The primal goal of this paper is to present an analytic solution via a co-state expression to the optimal control problem $(\mathcal{P})$.

We know from the Pontryagin principle [16] that if the control $\hat{u}$ is an optimal solution for the problem $(\mathcal{P})$, with $\hat{x}(.)$ and $\hat{\lambda}(.)$ denoting the state and co-state corresponding to $\hat{u}(.)$ respectively, then $\hat{u}$ is an extremal control, i.e. we have

$$\dot{\hat{x}} = H_\lambda(t, \hat{x}, \hat{u}, \hat{\lambda}) = A(t)\hat{x} + B(t)\hat{u}, \quad \hat{x}(0) = x_0; \tag{1.6}$$
$$\dot{\hat{\lambda}} = -H_x(t, \hat{x}, \hat{u}, \hat{\lambda}) = -A^T\hat{\lambda} - \nabla F(\hat{x}), \quad \hat{\lambda}(T) = \vec{0}. \tag{1.7}$$

and

$$H(t, \hat{x}(t), \hat{u}(t), \hat{\lambda}(t)) = \min_{\|u\|\leq 1} H(t, \hat{x}(t), u, \hat{\lambda}(t)), \quad a.e. \ t \in [0, T]. \tag{1.8}$$

By means of Pontryagin principle and the dynamic programming theory, many numerical algorithms have been suggested to approximate the solution of the problem $(\mathcal{P})$ [12]. This is due to the nonlinear integrand in the cost functional. It is even difficult for the case of $P(u)$ being nonconvex on the unit ball D in R^m. In this paper, as long as an optimal control of the problem $(\mathcal{P})$ exists, we solve the problem $(\mathcal{P})$ for an analytic solution by the co-state.

We see that, with respect to u, the minimization in (1.8) is equivalent to the following global optimization over a sphere:

$$\min_{\|u\|\leq 1} [P(u) + \hat{\lambda}(t)^T B(t)u], \quad a.e. \ t \in [0, T]. \tag{1.9}$$

When $P(u)$ is a nonconvex quadratic function, by the canonical dual transformation [7–9], the problem (1.9) can be solved completely. In [19], the global concave optimization over a sphere is solved by use of a differential system with the canonical dual function. Because the Pontryagin principle is a necessary condition for a control to be optimal, it is not sufficient for obtaining an optimal control to solve only the optimization (1.9). In this paper, combining the method given in [8,19] with the Pontryagin principle, also motivated by the significant works by C.Floudas et al. [1–3,6,10,13] on global optimizations, we solve the problem (1.1) and present the optimal control expressed by the co-state via canonical dual variables.

The rest of the paper is organized as follows. In Sect. 2, the backward differential flow is defined to deduce some optimality conditions for solving global optimizations by canonical dual functions. In Sect. 3, Some global optimization problems are solved by use of the backward differential flows. In the last section, the analytic solution to an optimal control problem is obtained via the expression of the co-state. Meanwhile, some examples are illustrated.

2 The backward differential flow and canonical dual function

Motivated by the significant work by Panos M. Pardalos and Vitaliy Yatsenko [15], in this section we present a differential flow for constructing the so called canonical dual function [8] to deal with the global optimization on $u \in R^m$

$$\min P(u),$$
$$\text{s.t.} u^T u \leq 1.$$

Here we use the method in our other paper (see [19]).

In the following we consider the function $P(u)$ to be twice continuously differentiable in R^m. Define the set

$$G = \left\{ \rho \geq 0 \mid [\nabla^2 P(u) + \rho I] > \vec{0}, u^T u \leq 1 \right\}. \tag{2.1}$$

By elementary calculus it is easy to get the following result.

Proposition 2.1 *G is an open set with respect to* $[0, +\infty)$*. If $\hat{\rho} \in G$, then $\rho \in G$ for $\forall \rho > \hat{\rho}$. If a $\rho^* \in G$ and a nonzero vector $u^* \in D = \{u^T u < 1\}$ satisfy following equation*

$$\nabla P(u^*) + \rho^* u^* = \vec{0}, \tag{2.2}$$

we focus on the flow $\hat{u}(\rho)$ which is well-defined near ρ^* by

$$\frac{d\hat{u}}{d\rho} + [\nabla^2 P(\hat{u}) + \rho I]^{-1} \hat{u} = \vec{0}, \tag{2.3}$$

$$\hat{u}(\rho^*) = u^*. \tag{2.4}$$

The flow $\hat{u}(\rho)$ can be extended to wherever $\rho \in G \cap [0, +\infty)$ [17]. The canonical dual function [8] with respect to a given flow $\hat{u}(\rho)$ is defined as follows:

$$P_d(\rho) = P(\hat{u}(\rho)) + \frac{\rho}{2} \hat{u}^T(\rho) \hat{u}(\rho) - \frac{\rho}{2}. \tag{2.5}$$

Lemma 2.1 *For a given flow defined by* (2.2)–(2.4)*, we have*

$$\frac{dP_d(\rho)}{d\rho} = \frac{1}{2} \hat{u}^T(\rho) \hat{u}(\rho) - \frac{1}{2}. \tag{2.6}$$

$$\frac{d^2 P_d(\hat{\rho})}{d\rho^2} = -\left(\frac{d\hat{u}(\rho)}{d\rho} \right)^T \left[\nabla^2 P(\hat{u}(\rho)) + \rho I \right] \frac{d\hat{u}(\rho)}{d\rho} \tag{2.7}$$

Proof Since $P_d(\rho)$ is differentiable,

$$\begin{aligned}
\frac{dP_d(\rho)}{d\rho} &= \frac{dP(\hat{u}(\rho))}{d\rho} + \frac{1}{2} \hat{u}^T(\rho) \hat{u}(\rho) + \frac{1}{2} \rho \frac{d(\hat{u}^T(\rho) \hat{u}(\rho))}{d\rho} - \frac{1}{2} \\
&= \nabla P(\hat{u}(\rho)) \frac{d(\hat{u}(\rho))}{d\rho} + \frac{1}{2} \hat{u}^T(\rho) \hat{u}(\rho) + \frac{1}{2} \rho \frac{d(\hat{u}^T(\rho) \hat{u}(\rho))}{d\rho} - \frac{1}{2}
\end{aligned}$$

$$= -\rho \hat{u}^T(\rho)\frac{d(\hat{u}(\rho))}{d\rho} + \frac{1}{2}\hat{u}^T(\rho)\hat{u}(\rho) + \frac{1}{2}\rho\frac{d(\hat{u}^T(\rho)\hat{u}(\rho))}{d\rho} - \frac{1}{2}$$
$$= \frac{-1}{2}\rho\frac{d(\hat{u}^T(\rho)\hat{u}(\rho))}{d\rho} + \frac{1}{2}\hat{u}^T(\rho)\hat{u}(\rho) + \frac{1}{2}\rho\frac{d(\hat{u}^T(\rho)\hat{u}(\rho))}{d\rho} - \frac{1}{2}$$
$$= \frac{1}{2}\hat{u}^T(\rho)\hat{u}(\rho) - \frac{1}{2}.$$

□

Further, since $P(x)$ is twice continuously differentiable, by (2.3) we have

$$\frac{d^2 P_d(\rho)}{d\rho^2} = \hat{u}^T(\rho)\frac{d\hat{u}(\rho)}{d\rho}$$
$$= -\left(\frac{d\hat{u}(\rho)}{d\rho}\right)^T \left[\nabla^2 P(\hat{u}(\rho)) + \rho I\right]\frac{d\hat{u}(\rho)}{d\rho}.$$

Lemma 2.2 *Let $\hat{u}(\rho)$ be a given flow defined by (2.2)–(2.4) and $P_d(\rho)$ be the corresponding canonical dual function defined by (2.5). We have* (i) *For every $\rho \in G$, $\frac{d^2 P_d(\rho)}{d\rho^2} \leq 0$;* (ii) *If $\hat{\rho} \in G$, then $\frac{dP_d(\rho)}{d\rho}$ monotonously decreases in $[\hat{\rho}, +\infty)$;* (iii) *If $\hat{\rho} \in G$ and $\hat{u}(\hat{\rho}) \in D$, in $[\hat{\rho}, +\infty)$, $P_d(\rho)$ is monotonously decreasing.*

Proof When $\rho \in G$, we have $\nabla^2 P(\hat{u}(\rho)) + \rho I > 0$. It follows from (2.7) that $\frac{d^2 P_d(\rho)}{d\rho^2} \leq 0$. Consequently, noting Proposition 2.1, we see that $\frac{dP_d(\rho)}{d\rho}$ monotonously decreases in $[\hat{\rho}, +\infty)$ when $\hat{\rho} \in G$. Finally, since $\hat{u}(\hat{\rho}) \in D$, $\frac{dP_d(\hat{\rho})}{d\rho} \leq 0$ by (2.6). It follows from $\hat{\rho} \in G$ that in $[\hat{\rho}, +\infty)$, $\frac{dP_d(\rho)}{d\rho} \leq 0$. Thus, in $[\hat{\rho}, +\infty)$, $P_d(\rho)$ is monotonously decreasing. □

Theorem 2.1 *If the flow $\hat{u}(\rho)$(defined by (2.2)–(2.4)) meets a boundary point of the ball $D = \{u \in R^m \mid \|u\| \leq 1\}$ at $\hat{\rho} \in G$, i.e.*

$$[\hat{u}(\hat{\rho})]^T \hat{u}(\hat{\rho}) = 1, \hat{\rho} \in G \tag{2.8}$$

then $\hat{u}$ is a global minimizer of $P(u)$ over the ball D. Further we have

$$\min_D P(u) = P(\hat{u}) = P_d(\hat{\rho}) = \max_{\rho \geq \hat{\rho}} P_d(\rho). \tag{2.9}$$

Proof By the definition of the flow $\hat{u}(\rho)$ ((2.2)–(2.4)) and Proposition 2.1, noting that $\hat{u}(\hat{\rho})$ is on the flow and $\hat{\rho} \in G$, we have, for all $\rho \geq \hat{\rho}$

$$\nabla\left\{P(\hat{u}(\rho)) + \frac{\rho}{2}\left[\hat{u}^T(\rho)\hat{u}(\rho) - 1\right]\right\} = \nabla P(\hat{u}(\rho)) + \rho\hat{u}(\rho) = \vec{0}, \tag{2.10}$$

and for all $\rho \geq \hat{\rho}$

$$\nabla^2(P(u) + \frac{\rho}{2}[u^T u - 1]) = \nabla^2 P(u) + \rho I > 0, \quad \forall u \in D. \tag{2.11}$$

In the following deducing, we need to note the fact that since $P(u)$ is twice continuously differentiable in R^n, there is a positive real δ such that (2.11) holds in $\{u^T u < 1 + \delta\}$ which contains D. In other words, for each $\rho > \hat{\rho}$, $\hat{u}(\rho)$ is the global minimizer of $P(u) + \frac{\rho}{2}[u^T u - 1]$ over D. Therefore, for every $u \in D = \{u \in R^n \mid u^T u \leq 1\}$, when $\rho \geq \hat{\rho}$, we have

$$P(u) \geq P(u) + \frac{\rho}{2}[u^T u - 1] \geq inf_D\left\{P(u) + \frac{\rho}{2}[u^T u - 1]\right\}$$
$$= P(\hat{u}(\rho)) + \frac{\rho}{2}\hat{u}^T(\rho)\hat{u}(\rho) - \frac{\rho}{2} = P_d(\rho). \tag{2.12}$$

Thus, by (2.10), Lemmas 2.2 and (2.8),

$$P(u) \geq max_{\rho \geq \hat{\rho}} P_d(\rho) = P_d(\hat{\rho}) = P(\hat{u}(\hat{\rho})) + \frac{\hat{\rho}}{2}[(\hat{u}(\hat{\rho}))^T \hat{u}(\hat{\rho}) - 1] = P(\hat{u}(\hat{\rho})) \quad (2.13)$$

Consequently

$$\min_D P(u) = \max_{\rho \geq \hat{\rho}} P_d(\rho). \quad (2.14)$$

This concludes the proof of Theorem 2.1. □

Similarly we have the following result.

Theorem 2.2 *Let the flow $\hat{u}(\rho)$ be defined by (2.2)–(2.4). If*

$$[\hat{u}(0)]^T \hat{u}(0) \leq 1, \quad (2.15)$$

then $\hat{u}(0)$ is a global minimizer of $P(u)$ over the ball D.

Proof By the definition of the flow $\hat{u}(\rho)$((2.2)–(2.4)), it follows from (2.16) that $\hat{\rho} = 0 \in G$ and $\hat{u}(0)$ is on the flow. We have, for all $\rho \geq 0$

$$\nabla \left\{ P(\hat{u}(\rho)) + \frac{\rho}{2}[\hat{u}^T(\rho)\hat{u}(\rho) - 1] \right\} = \nabla P(\hat{u}(\rho)) + \rho \hat{u}(\rho) = \vec{0},$$

and for all $\rho \geq 0$

$$\nabla^2 \left(P(u) + \frac{\rho}{2}[u^T u - 1] \right) = \nabla^2 P(u) + \rho I > 0, \ \ \forall u \in D.$$

Therefore, for every $u \in D = \{u \in R^n \mid u^T u \leq 1\}$, when $\rho \geq 0$, we have

$$\begin{aligned} P(u) \geq \ & P(u) + \frac{\rho}{2}[u^T u - 1] \geq inf_D \left\{ P(u) + \frac{\rho}{2}[u^T u - 1] \right\} \\ & = P(\hat{u}(\rho)) + \frac{\rho}{2}\hat{u}^T(\rho)\hat{u}(\rho) - \frac{\rho}{2} = P_d(\rho). \end{aligned}$$

Thus, noting (2.16) and using Lemma 2.2 ,

$$P(u) \geq \max_{\rho \geq 0} P_d(\rho) = P_d(0) = P(\hat{u}(0)) + \frac{0}{2}[(\hat{u}(0))^T \hat{u}(0) - 1] = P(\hat{u}(0)).$$

Consequently, $\hat{u}(0)$ is a global minimizer of $P(u)$ over the ball D. This concludes the proof of Theorem 2.1. □

Definition 2.1 Let $\hat{u}(\rho)$ be a flow defined by (2.2)–(2.4). We call $\hat{u}(\rho)$, $\rho \in (0, \rho^*]$ a backward differential flow.

In other words, the backward differential flow $\hat{u}(\rho)$, $\rho \in (0, \rho^*]$ comes from solving the Eq. (2.2) backwards from ρ^*.

In what follows, we introduce a result in [19] which is a sufficient condition for the global optimization over a sphere. Let $P(u)$ be strictly concave. Suppose that there are only finitely many of root pairs for (2.2) :

$$0 < \rho_1^* < \rho_2^* < \cdots < \rho_l^*,$$

associated with feasible points on the unit sphere:

$$\hat{u}_1, \hat{u}_2, ..., \hat{u}_l, \quad (2.16)$$

such that for each i,

$$\nabla P(\hat{u}_i) + \rho_i^* \hat{u}_i = 0, \quad \hat{u}_i^T \hat{u}_i = 1, \tag{2.17}$$

$$\rho_i^* > 0. \tag{2.18}$$

and for each $i = 1, 2, ..., l$, the associate state point $\hat{u}_i$ is uniquely corresponding to ρ_i^*.

Theorem 2.3 [19] *Let $P(u)$ be strictly concave. Suppose that (2.16)–(2.18) hold for the pairs $(\rho_i^*, \hat{u}_i)$, $i = 1, 2, ..., l$. If, for $i = 1, 2, ..., l$, $det[\nabla^2 P(\hat{u}_i) + \rho_i^* I] \neq 0$ and $\frac{d^2 P_d(\rho_i^*)}{d\rho^{*2}} > 0$, then $\hat{u}_l$ is the unique global minimizer of (1.1).*

Before going further let us give the following remark:

Remark 2.4 In this section, the idea to introduce the set G of shift parameters is closely following the works by Floudas et al. [1–3]. In [2], they developed a global optimization method, αBB, for general twice-differentiable constrained optimizations. It is a powerful theory which can be used to solve a very broad class of global optimization problems. In [1], the performance of the proposed algorithm [2] and its alternative under-estimators is studied through their application to a variety of problems. In this paper we use the method given by Floudas et al. [2] proposing to utilize some α parameter to generate valid convex under-estimators for nonconvex terms of generic structure.

In the following two sections we will use Theorem 2.1–2.3 to solve the optimization problems and optimal control problem ($\mathcal{P}$).

3 Find the global minimizer by the backward differential flow

Here we propose using backward differential flow with the corresponding canonical dual function to solve the optimization problem $\min_D P(u)$. We focus on Theorems 2.1 and 2.3. The main idea of using backward differential flows to find global minimizer is as follows. Since D is bounded and $P(u)$ is twice continuously differentiable, we can choose a large positive parameter ρ^* such that $\nabla^2 P(u) + \rho^* I > 0, \ \ \forall u \in D$ and $\rho^* > \sup_D \{\|\nabla P(u)\|\}$. If $\nabla P(0) \neq 0$, then it follows from $\|\frac{\nabla^2 P(u)}{\rho^*}\| < 1$ uniformly in D that there is a unique nonzero fixed point $u^* \in D$ such that

$$\frac{-\nabla P(u^*)}{\rho^*} = u^* \tag{3.1}$$

by Brown fixed-point theorem. It means the pair (u^*, ρ^*) satisfies (2.2). We solve (2.2) backwards from ρ^* to get the backward flow $\hat{u}(\rho)$, $\rho \in [0, \rho^*]$. If there is a $\hat{\rho} \in G \cap (0, \rho^*]$ such that $\hat{u}(\hat{\rho})^T \hat{u}(\hat{\rho}) = 1$, then by Theorem 2.1 we see that $\hat{u}(\hat{\rho})$ is a global minimizer of $P(u)$ over D. On the other hand, if $\hat{u}(0)^T \hat{u}(0) < 1$, then by Theorem 2.2 we see that $\hat{u}(0)$ is a global minimizer of $P(u)$ over D.

On how to choose the desired parameter ρ^*, we may be referred to [1,2]. Some good algorithms are given in [1,2] to estimate the bounds of $\|\nabla^2 P(u)\|$. If there is a positive real number M such that $\|\nabla^2 P(u)\| \leq M, \ \ \forall u \in D$, then a properly large parameter ρ^* can be obtained by the inequalities

$$\frac{\|\nabla^2 P(u)\|}{\rho^*} \leq \frac{M}{\rho^*} < 1, \quad \rho^* > \sup_D \{\|\nabla P(u)\|\}$$

uniformly on D for us to use Brown fixed-point theorem. We should choose

$$\rho^* > \max\{\sup_D \|\nabla^2 P(u)\|,\ \sup_D\{\|\nabla P(u)\|\}.$$

We will discuss to calculate the parameter ρ^* in detail by use of the results in [3,6,13] with the future works.

In the following we present several examples to find global minimizers by backward differential flows.

Example 3.1 (A concave minimization). Let us consider the following one dimensional concave minimization problem

$$p^* = \min P(u) = \frac{-1}{12}u^4 - u^2 + u,$$
$$\text{s.t. } u^2 \leq 1.$$

We have $P'(u) = \frac{-1}{3}u^3 - 2u + 1,\quad P''(u) = -u^2 - 2 < 0, \forall u^2 \leq 1$. Choosing $\rho^* = 10$, solve the following equation in $\{u^2 < 1\}$(for the fixed point)

$$\frac{-1}{3}u^3 - 2u + 1 + 10u = 0$$

to get $u^* = -0.1251$. Next we solve the following backward differential equation

$$\frac{du(\rho)}{d\rho} = \frac{u(\rho)}{u^2(\rho) + 2 - \rho},\quad u(\rho^*) = -0.1251,\quad \rho \leq 10.$$

To find a parameter such that

$$u^2(\rho) = 1,$$

we get

$$\hat{\rho} = \frac{10}{3},$$

which satisfies

$$P''(u) + \frac{10}{3} > 0, \forall u^2 \leq 1.$$

Let $u\left(\frac{10}{3}\right)$ be denoted by $\hat{u}$. Compute the solution of following algebra equation

$$\frac{-1}{3}u^3 - 2u + 1 + \frac{10}{3}u = 0,\quad u^2 = 1$$

to get $\hat{u} = -1$. It follows from Theorem 2.1 that $\hat{u} = -1$ is the global minimizer of $P(u)$ over $[-1, 1]$.

Remark 3.1 For the proper parameter ρ^*, it is worth investigating how to get the solution u^* of the Eq. (3.1) inside of D. For this issue, when $P(u)$ is a polynomial we may be referred to [10]. There are results in [10] on bounding the zeros of a polynomial. We may consider for a given bounds to determine the parameter by use of the results in [10] on the relation between the zeros and the coefficients. We will discuss it with the future works as well.

Example 3.2 (A non-convex quadratic optimization over the sphere) . Given a symmetric matrix $G \in R^{m\times m}$ and a vector $f \in R^m$, $f \neq 0$, let $P(u) = \frac{1}{2}u^T Gu - f^T u$ be non-convex. We consider the following global optimization over a sphere

$$\min P(u) = \frac{1}{2}u^T Gu - f^T u,$$
$$\text{s.t.} u^T u \leq 1.$$

Suppose that G has $p \leq m$ distinct eigenvalues $a_1 < a_2 < \cdots < a_p$. Since $P(x) = \frac{1}{2}u^T Gu - f^T u$ is non-convex, $a_1 < 0$. Choose a large $\rho^* > -a_1$ such that

$$0 < \|(G + \rho^* I)^{-1} f\| < 1,$$

noting that $f \neq 0$. We see that the backward differential equation is

$$\frac{du}{d\rho} = -(G + \rho I)^{-1}u, \;\; u(\rho^*) = (G + \rho^* I)^{-1} f, \;\; \rho \leq \rho^*,$$

which leads a backward flow

$$u(\rho) = (G + \rho I)^{-1} f, \;\; \rho \leq \rho^*.$$

Further noting that there is an orthogonal matrix R leading to a diagonal transformation $RGR^T = D := (a_i\delta_{ij})$ and correspondingly $Rf = g := (g_i)$, we have

$$u^T(\rho)u(\rho) = f^T(G + \rho I)^{-2} f = \sum_{i=1}^{p} \frac{g_i^2}{(a_i + \rho)^2}.$$

Since $f^T(G + \rho^* I)^{-2} f < 1$ and

$$\lim_{\rho > -a_1, \rho \to -a_1} \sum_{i=1}^{p} \frac{g_i^2}{(a_i + \rho)^2} = +\infty,$$

there is the unique $\hat{\rho} : -a_1 < \hat{\rho} < \rho^*$ such that

$$u^T(\hat{\rho})u(\hat{\rho}) = f^T(G + \hat{\rho} I)^{-2} f = \sum_{i=1}^{p} \frac{g_i^2}{(a_i + \hat{\rho})^2} = 1.$$

By Theorem 2.1, we see that $u(\hat{\rho}) = (G + \hat{\rho} I)^{-1} f$ is a global minimizer of the problem.

Example 3.3 (A convex quadratic optimization over the sphere) . Given a positive definite matrix $G \in R^{m\times m}$ and a vector $f \in R^m$, $f \neq 0$, let $P(u) = \frac{1}{2}u^T Gu - f^T u$. We consider the following convex optimization over a sphere

$$\min P(u) = \frac{1}{2}u^T Gu - f^T u,$$
$$\text{s.t.} \;\; u^T u \leq 1.$$

Choose a large $\rho^* > 0$ such that

$$0 < \|(G + \rho^* I)^{-1} f\| < 1,$$

noting that $f \neq 0$. We see that the backward differential equation is

$$\frac{du}{d\rho} = -(G + \rho I)^{-1}u, \;\; u(\rho^*) = (G + \rho^* I)^{-1} f, \;\; \rho \leq \rho^*,$$

which leads a backward flow

$$u(\rho) = (G + \rho I)^{-1} f, \quad \rho \le \rho^*.$$

Since $G > 0$ and $f \neq 0$, by Lemma 2.1, we have, for $\rho \in [0, \rho^*]$,

$$\frac{d(u^T u)}{d\rho} = -2 f^T [G + \rho I]^{-3} f < 0. \tag{3.2}$$

If $f^T G^{-2} f \le 1$, then by Theorem 2.2 we see that $u(0) = G^{-1} f$ is the global minimizer of $P(u)$ over the unit ball D.

On the other hand, we consider the case $f^T G^{-2} f > 1$, noting (3.2), there is the unique $\hat{\rho} \in [0, \rho^*)$ such that $u(\hat{\rho})^T u(\hat{\rho}) = 1$. Actually, in this case, we can solve

$$f^T (G + \rho I)^{-2} f = 1, \quad \rho \in [0, \rho^*] \tag{3.3}$$

to get this unique parameter $\hat{\rho}$. By use of Theorem 2.1, we see that $u(\hat{\rho})) = (G + \hat{\rho} I)^{-1} f$ is the global minimizer of $P(u)$ over the unit ball D.

4 The analytic solution of an optimal control problem

In this section, we consider $A(t)$, $B(t)$ in the problem (1.1) to be constant matrices, $F(x) = c^T x$ and

$$P(u) = \frac{1}{2} u^T G u - b^T u,$$

where $c \in R^{n \times 1}$, $b \in R^{m \times 1}$ and $G \in R^{m \times m}$, is symmetric. Suppose that G has $p \le m$ distinct eigenvalues $a_1 < a_2 < \cdots < a_p (p \ge 1)$.

We need the following **basic assumption:**

$$\operatorname{rank}(B^T, b) > \operatorname{rank}(B^T). \tag{*}$$

We consider the following optimal control problem:

$$(\mathcal{P}) \quad \min J(u) = \int_0^T \left[c^T x + \tfrac{1}{2} u^T G u - b^T u \right] dt, \tag{4.1}$$

$$\text{s.t. } \dot{x} = Ax + Bu, \quad x(0) = x_0, \quad t \in [0, T], \quad \|u\| \le 1. \tag{4.2}$$

To solve the above problem, we define the function $\phi(t, x) = \psi^T(t) x$, where the continuously differentiable function $\psi(t)$ is to be determined by the following ordinary differential equation:

$$\dot{\psi}(t) = -A^T \psi(t) + c, \tag{4.3}$$

$$\psi(T) = \vec{0}. \tag{4.4}$$

Compare (4.3)–(4.4) with (1.7), we see that

$$\psi(t) = -\lambda(t), \quad a.e.\ t \in [0, T]. \tag{4.5}$$

We have

$$\begin{aligned}
J(u) &= \int_0^T \left[c^T x + \frac{1}{2}u^T Gu - b^T u\right] dt \\
&= \int_0^T \left[(\dot{\psi}(t) + A^T\psi(t))^T x + \frac{1}{2}u^T Gu - b^T u\right] dt \\
&= \int_0^T \left[\dot{\psi}^T(t)x + \psi(t)^T Ax + \frac{1}{2}u^T Gu - b^T u\right] dt \\
&= \int_0^T \left[\dot{\psi}^T(t)x + \psi(t)^T(Ax + Bu) - \psi(t)^T Bu + \frac{1}{2}u^T Gu - b^T u\right] dt \\
&= \int_0^T \left[\dot{\psi}^T(t)x(t) + \psi(t)^T\dot{x}(t) - \psi(t)^T Bu + \frac{1}{2}u^T Gu - b^T u\right] dt \\
&= \int_0^T \left[\dot{\phi}(t, x(t)) - \psi(t)^T Bu + \frac{1}{2}u^T Gu - b^T u\right] dt \\
&= \phi(T, x(T)) - \phi(0, x(0)) + \int_0^T \left[\frac{1}{2}u^T Gu - b^T u - \psi(t)^T Bu\right] dt \\
&= -\phi(0, x(0)) + \int_0^T \left[\frac{1}{2}u^T Gu - b^T u - \psi(t)^T Bu\right] dt, \qquad (4.6)
\end{aligned}$$

noting that $\psi(T) = \vec{0}$ and $x(0) = x_0$. Thus,

$$\min J(u) = -\phi(0, x(0)) + \min \int_0^T \left[\frac{1}{2}u^T Gu - b^T u - \psi(t)^T Bu\right] dt. \qquad (4.7)$$

Consequently, we deduce that, for $a.e.$ $t \in [0, T]$, the optimal control

$$\hat{u}(t) = \text{argmin} \left\{\frac{1}{2}u^T Gu - b^T u - \psi(t)^T Bu \mid u^T u \leq 1\right\}. \qquad (4.8)$$

By the relation between $\psi(t)$ and the co-state in (4.5), for each $t \in [0, T]$, we need to solve following non-convex optimization

$$\min \tfrac{1}{2}u^T Gu - (b - B^T\lambda(t))^T u, \qquad (4.9)$$

$$\text{s.t.} u^T u \leq 1. \qquad (4.10)$$

It follows from the basic assumption (*) rank$(B^T, b) >$ rank(B^T) that $b - B^T\lambda(t) \neq 0$ for each $t \in [0, T]$. By Example 3.2, for each $t \in [0, T]$, we have

$$\hat{u}(t) = (G + \rho_t I)^{-1}[b - B^T\lambda(t)] \qquad (4.11)$$

where the dual variable $\rho_t > -a_1$ satisfies

$$(b - B^T\lambda(t))^T(G + \rho_t I)^{-2}(b - B^T\lambda(t)) = 1. \quad (4.12)$$

With respect to λ we define the function $\rho(\lambda)$ by the following equation

$$(b - B^T\lambda)^T(G + \rho(\lambda)I)^{-2}(b - B^T\lambda) = 1, \quad (4.13)$$

$$\rho(\lambda) > -a_1. \quad (4.14)$$

We have the optimal control by the co-state

$$\hat{u} = (G + \rho(\lambda)I)^{-1}(b - B^T\lambda). \quad (4.15)$$

On the other hand, by the solution of the ordinary differential Eqs. (4.3)–(4.4), we have

$$\lambda(t) = -\psi(t) = e^{A^T T}\int_0^{T-t} e^{-A^T t}e^{-A^T s}dsc = e^{A^T(T-t)}\left[\int_0^{T-t} e^{-A^T s}ds\right]c. \quad (4.16)$$

Example 4.1 Consider $G = -I$ in the problem (4.1)–(4.2), i.e.

$$P(u) = \frac{-1}{2}u^T u - b^T u.$$

By the Pontryagin principle, we need to solve a system on the state and co-state

$$\dot{\hat{x}} = A\hat{x} + B\hat{u}, \ \ \hat{x}(0) = x_0, \quad (4.17)$$

$$\dot{\hat{\lambda}} = -A^T\hat{\lambda} - c, \ \ \hat{\lambda}(T) = \vec{0}, \quad (4.18)$$

and a global concave optimization for almost each $t \in [0, T]$,

$$\min_{\|u\|\leq 1}\left[\frac{-1}{2}u^T u - b^T u + \hat{\lambda}^T(t)Bu\right]. \quad (4.19)$$

By the basic assumption (*), we have $b - B^T\hat{\lambda}(t) \neq 0, \forall t \in [0, T]$. By Example 3.2, we have

$$\hat{u}(t) = (\hat{\rho} - 1)^{-1}(b - B^T\hat{\lambda}(t)),$$

where $\hat{\rho}$ is determined by the following equation:

$$(\rho - 1)^{-2}(b - B^T\hat{\lambda})^T(b - B^T\hat{\lambda}) = 1, \ \ \rho > 0, \ \rho \neq 1.$$

It gives two possible choices of roots

$$\rho = 1 \pm \|b - B^T\hat{\lambda}\|. \quad (4.20)$$

By Theorem 2.1 and Example 3.2(or directly by observing (4.19)), we have to take

$$\hat{\rho} = 1 + \|b - B^T\hat{\lambda}\|. \quad (4.21)$$

We see that at each $t \in [0, T]$, the global concave optimization (4.19) has the unique solution

$$\hat{u}(t) = (\hat{\rho} - 1)^{-1}(b - B^T\hat{\lambda}(t)) = \frac{b - B^T\hat{\lambda}(t)}{\|b - B^T\hat{\lambda}(t)\|}. \quad (4.22)$$

It follows from the traditional optimal control theory that $\hat{u}(t)$ is the optimal control which is an analytic solution expressed by the co-state.

Example 4.2 Consider the following problem

$$(\mathcal{P}) \quad \min \int_0^1 [x + (-u^4 - u^2 + u)]dt, \tag{4.23}$$

$$\text{s.t. } \dot{x} = x + u, \quad x(0) = 0, \quad t \in [0, 1], \quad |u| \leq 1.$$

In this example we have

$$H(x, u, \lambda) = \lambda(x + u) + x - u^4 - u^2 + u.$$

The adjoint equation

$$\dot{\lambda} = -\lambda - 1, \quad \lambda(T) = 0$$

gives

$$\lambda(t) = \int_0^{1-t} e^{1-t-s} ds = e^{1-t} - 1.$$

For each $t \in [0, 1]$ such that $1 + \lambda(t) \neq 0$, we solve

$$\min_{|u| \leq 1} \{-u^4 - u^2 + u(1 + \lambda(t))\}. \tag{4.24}$$

By Canonical dual method, we have

$$-4u^3 - 2u + (1 + \lambda(t)) + \rho u = 0, \quad u^2 = 1,$$

and

$$(\rho - 6)u = -(1 + \lambda), \quad 1 + \lambda \neq 0.$$

As in Example 3.2 of [19], using Theorem 2.3, we get $\hat{\rho} = 6 + |1 + \lambda|$ and obtain the unique global minimizer of (4.24)

$$\hat{u}(t) = \frac{-(1 + \lambda)}{|1 + \lambda|}, \quad 1 + \lambda \neq 0.$$

Since

$$1 + \lambda(t) = e^{1-t} > 0, \quad \forall t \in [0, 1],$$

we get the optimal control for the problem (4.23)

$$\hat{u}(t) \equiv -1, \quad \forall t \in [0, 1].$$

Remark We can also use the process (4.6)–(4.8) to show that $\hat{u}(t) = \frac{-(1+\lambda)}{|1+\lambda|}$ is the optimal control.

References

1. Adjiman, C.S., Androulakis, I.P., Floudas, C.A.: A global optimization method, αBB, for general twice-differentiable constraied NLP's—II, implementation and computational results. Comput. Chem. Eng. **22**(9), 1159–1179 (1998)
2. Adjiman, C.S., Dallwig, S., Floudas, C.A., Neumaier, A.: A global optimization method, αBB, for general twice-differentiable constraied NLP's—I, theoretical advances. Comput. Chem. Eng. **22**(9), 1137–1158 (1998)

3. Adjiman, C.S., Androulakis, I.P., Floudas, C.A.: Global optimization of mixed-integer nonlinear problems. AIChE. J. **46**(9), 1769–1797 (2000)
4. Butenko, S., Murphey, R., Pardalos, P.M. (eds.): Recent Developments in Cooperative Control and Optimization. Kluwer, Dordrecht (2003)
5. Floudas, C.A., Visweswaran, V.: Quadratic Optimization. In: Horst, R., Pardalos, P.M. (eds.) Handbook of Global Optimization, pp. 217–270. Kluwer, Dordrecht (1995)
6. Floudas, C.A. (ed.): Deterministic Global Optimization: Theory, Methods and Applications. Kluwer, Dordrecht (2000)
7. Gao, D.Y.: Duality Principles in Nonconvex Systems, Theory, Methods and Applications. Kluwer, Dordrecht (2000)
8. Gao, D.Y.: Canonical duality theory and solutions to constrained nonconvex quadratic programming. J. Glob. Optim. **29**, 377–399 (2004)
9. Gao, D.Y.: Solutions and optimality criteria to box constrained nonconvex minimization problems. J. Ind. Manage. Optim. **3**(2), 293–304 (2007)
10. Hertz, D., Adjiman, C.S., Floudas, C.A.: Two results on bounding the roots of interval polynomials. Comput. Chem. Eng. **23**, 1333–1339 (1999)
11. Hirsch, M.J., Commander, C., Pardalos, P.M., Murphey, R. (eds.): Optimization and Cooperative Control Strategies, Lecture Notes in Control and Information Sciences, Vol. 381. Springer, New York (2009)
12. Jinghao, Z., Zhiqiang, Z.: An approximation to discrete optimal controls. Int. J. Math. Math. Sci. **47**, 2989–3001 (2003)
13. Monnigmann M.: Efficient calculation of bounds on spectra of Hessian matrices. SIAM J. Sci. Comput. **30**(5), 2340–2357
14. Pardalos, P.M., Tseveendorj, I., Enkhbat, R. (eds.): Optimization and Optimal Control. World Scientific, Singapore (2003)
15. Pardalos, P.M., Yatsenko, V.A.: Optimization and Control of Bilinear Systems. Springer, New York (2008)
16. Russell, D.L.: Mathematics of Finite Dimensional Control Systems: Theory and Design. Marcel Dekker, New York (1979)
17. Robinson, C.: Dynamical systems. CRC Press, London (1995)
18. Sontag, E.D.: Mathematical Control Theory: Deterministic Finite Dimensional Systems, (2nd edn.). Springer, New York (1998)
19. Zhu, J., Tao, S., Gao, D.: A study on concave optimization via canonical dual function. J. Comput. Appl. Math. **224**, 459–464 (2009)