

Research Article

Complete Solutions to General Box-Constrained Global Optimization Problems

Dan Wu and Youlin Shang

Department of Mathematics, Henan University of Science and Technology, Luoyang 471003, China

Correspondence should be addressed to Dan Wu, lywd2964@126.com

Received 17 March 2011; Accepted 16 May 2011

Academic Editor: Juan Manuel Peña

Copyright © 2011 D. Wu and Y. Shang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a global optimization method for solving general nonlinear programming problems subjected to box constraints. Regardless of convexity or nonconvexity, by introducing a differential flow on the dual feasible space, a set of complete solutions to the original problem is obtained, and criteria for global optimality and existence of solutions are given. Our theorems improve and generalize recent known results in the canonical duality theory. Applications to a class of constrained optimal control problems are discussed. Particularly, an analytical form of the optimal control is expressed. Some examples are included to illustrate this new approach.

1. Introduction

In this paper, we consider the following general box constrained nonlinear programming problem (the primal problem (\mathcal{P}) in short):

$$(\mathcal{P}) : \min\{P(x) \mid x \in \mathcal{X}_a\}, \quad (1.1)$$

where $\mathcal{X}_a = \{x \in \mathcal{R}^n \mid \ell^l \leq x \leq \ell^u\}$ is a feasible space, $\ell^l, \ell^u \in \mathcal{R}^n$ are two given vectors, and $P(x)$ is twice continuously differentiable in \mathcal{R}^n . Here, we discuss the primal problem (\mathcal{P}) involving nonconvexity or convexity in the objective function.

Problem (1.1) appears in many applications, such as engineering design, phase transitions, chaotic dynamics, information theory, and network communication [1, 2]. Particularly, if $\ell^l = \{0\}$ and $\ell^u = \{1\}$, the problem leads to one of the fundamental problems in combinatorial optimization, namely, the integer programming problem [3]. By the fact that the feasible space \mathcal{X}_a is a closed convex subset of \mathcal{R}^n , the primal problem has at least one global minimizer. When (\mathcal{P}) is a convex programming problem, a global minimizer can be

obtained by many well-developed nonlinear optimization methods based on the Karush-Kuhn-Tucker (or simply KKT) optimality theory [4]. However, for (\mathcal{D}) with nonconvexity in the objective function, traditional KKT theory and direct methods can only be used for solving (\mathcal{D}) to local optimality. So, our interest will be mainly in the case of $P(x)$ being nonconvex on \mathcal{X}_a in this paper. For special cases of minimizing a nonconvex quadratic function subject to box constraints, much effort and progress have been made on locating the global optimal solution based on the canonical duality theory by Gao (see [5–7] for details). As indicated in [8], the key step of the canonical duality theory is to introduce a canonical dual function, but commonly used methods are not guaranteed to construct it since the general form of the objective function given in (1.1). Thus, there has been comparatively little work in global optimality for general cases.

Inspired and motivated by these facts, a differential flow for constructing the canonical dual function is introduced and a new approach to solve the general (especially nonconvex) nonlinear programming problem (\mathcal{D}) is investigated in this paper. By means of the canonical dual problem, some conditions in global optimality are deduced, and global and local extrema of the primal problem can be identified. An application to the linear-quadratic optimal control problem with constraints is discussed. These results presented in this paper can be viewed as an extension and an improvement in the canonical duality theory [8–10].

The paper is organized as follows. In Section 2, a differential flow is introduced to present a general form of the canonical dual problem to (\mathcal{D}) . The relation of this transformation with the classical Lagrangian method is discussed. In Section 3, we present a set of complete solutions to (\mathcal{D}) by the way presented in Section 2. The existence of the canonical dual solutions is also given. We give an analytic solution to the box-constrained optimal control problem via canonical dual variables in Section 4. Meanwhile, some examples are used to illustrate our theory.

2. A Differential Flow and Canonical Dual Problem

In the beginning of this paper, we have mentioned that our primal goal is to find the global minimizers to a general (mainly nonconvex) box-constrained optimization problem (\mathcal{D}) . Due to the assumed nonconvexity of the objective function, the classical Lagrangian $L(x, \sigma)$ is no longer a saddle function, and the Fenchel-Young inequality leads to only a weak duality relation: $\min P \geq \max P^*$. The nonzero value $\theta = \min P(x) - \max P^*(\sigma)$ is called the *duality gap*, where possibly, $\theta = \infty$. This duality gap shows that the well-developed Fenchel-Moreau-Rockafellar duality theory can be used mainly for solving convex problems. Also, due to the nonconvexity of the objective function, the problem may have multiple local solutions. The identification of a global minimizer has been a fundamentally challenging task in global optimization. In order to eliminate this duality gap inherent in the classical Lagrange duality theory, a so-called *canonical duality theory* has been developed [2, 9]. The main idea of this new theory is to introduce a *canonical dual transformation* which may convert some nonconvex and/or nonsmooth primal problems into smooth *canonical dual problems* without generating any duality gap and deduce some global solutions. The key step in the canonical dual transformation is to choose the (nonlinear) geometrical operator $\Lambda(x)$. Different forms of $\Lambda(x)$ may lead to different (but equivalent) canonical dual functions and canonical dual problems. So far, in most literatures related, the canonical dual transformation is discussed and the canonical dual function is formulated in quadratic minimization problems (i.e., the objective function is the quadratic form). However, for the general form of the objective

function given in (1.1), in general, it lacks effective strategies to get the canonical dual function (or the canonical dual problem) by commonly used methods. The novelty of this paper is to introduce the *differential flow* created by differential equation (2.6) to construct the canonical dual function for the problem (\mathcal{P}) . Lemma 2.5 guarantees the existence of the differential flow; Theorem 2.3 shows that there is no duality gap between the primal problem (\mathcal{P}) and its canonical dual problem (\mathcal{P}^d) given in (2.7) via the differential flow; Meanwhile, Theorems 3.1–3.4 use the differential flow to present a global minimizer. In addition, the idea to introduce the set \mathcal{S} of shift parameters is closely following the works by Floudas et al. [11, 12]. In [12], they developed a global optimization method, *αBB*, for general twice-differentiable constrained optimizations proposing to utilize some α parameter to generate valid convex under estimators for nonconvex terms of generic structure.

The main idea of constructing the differential flow and the canonical dual problem is as follows. For simplicity without generality, we assume that $\ell^u = -\ell^l = -\ell^{1/2} = \{\sqrt{\ell_i}\}$, namely, $\mathcal{X}_a = \{x \in \mathcal{R}^n \mid -\sqrt{\ell_i} \leq x_i \leq \sqrt{\ell_i}\}$, where $\ell_i \neq 0$ for all i .

Let \mathcal{S} denote the dual feasible space

$$\mathcal{S} = \left\{ \rho \in \mathcal{R}_+^n \mid \left[\nabla^2 P(x) + \text{Diag}(\rho) \right] > 0, \forall x \in \mathcal{X}_a \right\}, \quad (2.1)$$

where $\mathcal{R}_+^n := \{\rho \in \mathcal{R}^n \mid \rho \geq 0\}$, and $\text{Diag}(\rho) \in \mathcal{R}^{n \times n}$ is a diagonal matrix with $\rho_i, i = 1, 2, \dots, n$, as its diagonal entries.

Lemma 2.1. *The dual feasible space \mathcal{S} is an open convex subset of \mathcal{R}_+^n . If $\hat{\rho} \in \mathcal{S}$, then $\rho \in \mathcal{S}$ for any $\rho \geq \hat{\rho}$.*

Proof. Notice that $P(x)$ is twice continuously differentiable in \mathcal{R}^n . For any $x \in \mathcal{X}_a$, the Hessian matrix $\nabla^2 P(x)$ is a symmetrical matrix. We know that for any given $Q = Q^T \in \mathcal{R}^{n \times n}$, $\{\rho \in \mathcal{R}_+^n \mid Q + \text{Diag}(\rho) > 0\}$ is a convex set. By the fact that the intersection of any collection of convex sets is convex, the dual feasible space \mathcal{S} is an open convex subset of \mathcal{R}_+^n . In addition, it follows from the definition of \mathcal{S} that $\rho \in \mathcal{S}$ for any $\rho \geq \hat{\rho}$. This completes the proof. \square

Suppose that $\rho^* \in \mathcal{S}$ and a nonzero vector $x^* \in \mathcal{X}_a$ satisfy

$$\nabla P(x^*) + \text{Diag}(\rho^*)x^* = 0. \quad (2.2)$$

A differential flow $x(\rho)$ is introduced over a relative small neighborhood of ρ^* such that

$$\nabla P(x(\rho)) + \text{Diag}(\rho)x(\rho) = 0, \quad x(\rho^*) = x^*, \quad (2.3)$$

which is equivalent to

$$\nabla^2 P(x(\rho))\nabla x(\rho) + \text{Diag}(\rho)\nabla x(\rho) + \text{Diag}(x(\rho)) = 0, \quad x(\rho^*) = x^*, \quad (2.4)$$

where $\nabla x(\rho)$ is the Jacobian of x and is a matrix whose ij th entry is equal to the partial derivative $\partial x_i / \partial \rho_j$. Here, we hope to preserve invertibility of the matrix $\nabla^2 P(x(\rho)) + \text{Diag}(\rho)$ on the choice of the neighborhood of ρ^* .

Let

$$H(\rho, x(\rho)) = -\left[\nabla^2 P(x(\rho)) + \text{Diag}(\rho)\right]^{-1}. \quad (2.5)$$

Then, a differential flow $x(\rho)$ can be defined by the following differential system:

$$\begin{aligned} dx_i(\rho) &= H_{i1}(\rho)x_1(\rho)d\rho_1 + H_{i2}(\rho)x_2(\rho)d\rho_2 + \cdots + H_{in}(\rho)x_n(\rho)d\rho_n, \\ x_i(\rho^*) &= x_i^*, \quad i = 1, 2, \dots, n, \end{aligned} \quad (2.6)$$

where $H_{ij}(\rho)$ is the ij th entry of $H(\rho, x(\rho))$. Based on the Extension theory [13, 14], the solution $x(\rho)$ of the differential system (2.6) can be extended to a space in \mathcal{S} . The canonical dual function is defined as

$$P^d(\rho) = P(x(\rho)) + \sum_{i=1}^n \frac{\rho_i}{2} [x_i^2(\rho) - \ell_i]. \quad (2.7)$$

Thus, the canonical dual problem for our primal problem (\mathcal{P}) can be proposed as follows

$$(P^d) : \max \left\{ P^d(\rho) = P(x(\rho)) + \sum_{i=1}^n \frac{\rho_i}{2} [x_i^2(\rho) - \ell_i] \mid \rho \in \mathcal{S} \right\}. \quad (2.8)$$

In the following, we show that (P^d) is canonically (i.e., with zero duality gap) dual to (\mathcal{P}) .

Lemma 2.2. *Let $x(\rho)$ be a given flow defined by (2.6), and $P^d(\rho)$ be the corresponding canonical dual function defined by (2.7). For any $\rho \in \mathcal{S}$, we have*

$$\begin{aligned} \nabla P^d(\rho) &= \left(\frac{1}{2} [x_1^2(\rho) - \ell_1], \frac{1}{2} [x_2^2(\rho) - \ell_2], \dots, \frac{1}{2} [x_n^2(\rho) - \ell_n] \right)^T, \\ \nabla^2 P^d(\rho) &= -\text{Diag}(x(\rho)) \left[\nabla^2 P(x(\rho)) + \text{Diag}(\rho) \right]^{-1} \text{Diag}(x(\rho)). \end{aligned} \quad (2.9)$$

Proof. Since $P^d(\rho)$ is differentiable, for any $\rho \in \mathcal{S}$,

$$\begin{aligned} \frac{\partial P^d(\rho)}{\partial \rho_i} &= \nabla P(x(\rho))^T \frac{\partial x(\rho)}{\partial \rho_i} + \frac{1}{2} [x_i^2(\rho) - \ell_i] + \sum_{j=1}^n \rho_j x_j(\rho) \frac{\partial x_j(\rho)}{\partial \rho_i} \\ &= \nabla P(x(\rho))^T \frac{\partial x(\rho)}{\partial \rho_i} + \rho^T \text{Diag}(x(\rho)) \frac{\partial x(\rho)}{\partial \rho_i} + \frac{1}{2} [x_i^2(\rho) - \ell_i] \\ &= (\nabla P(x(\rho)) + \text{Diag}(\rho)x(\rho))^T \frac{\partial x(\rho)}{\partial \rho_i} + \frac{1}{2} [x_i^2(\rho) - \ell_i]. \end{aligned} \quad (2.10)$$

It follows from the process (2.2)–(2.6) that

$$\frac{\partial P^d(\rho)}{\partial \rho_i} = \frac{1}{2} [x_i^2(\rho) - \ell_i]. \quad (2.11)$$

From (2.6), we have $\partial x_i(\rho) / \partial \rho_j = H_{ij}(\rho) x_j(\rho)$. By (2.11), then

$$\frac{\partial^2 P^d(\rho)}{\partial \rho_i \partial \rho_j} = x_i(\rho) \frac{\partial x_i(\rho)}{\partial \rho_j} = x_i(\rho) H_{ij}(\rho) x_j(\rho). \quad (2.12)$$

By the definition of $H(\rho, x(\rho))$, this concludes the proof of Lemma 2.2. \square

By Lemma 2.2, the canonical dual function $P^d(\rho)$ is concave on \mathcal{S} . For a critical point, $\hat{\rho} \in \mathcal{S}$, $\hat{\rho}$ must be a global maximizer of (P^d) , and it can be solved by many well-developed nonlinear programming methods. If $\hat{\rho} \in \mathcal{S}$ and $x(\hat{\rho}) \in \mathcal{X}_a$, we have $\partial P^d(\hat{\rho}) / \partial \hat{\rho}_i = (1/2)[x_i^2(\hat{\rho}) - \ell_i] \leq 0$, and for any $\rho \geq \hat{\rho}$, by negative definiteness of $H(x(\rho), \rho)$,

$$\frac{\partial [(1/2)(x_i^2(\rho) - \ell_i)]}{\partial \rho_i} = x_i(\rho) \frac{\partial x_i(\rho)}{\partial \rho_i} = H_{ii}(\rho) x_i^2(\rho) \leq 0. \quad (2.13)$$

Thus for any $\rho \geq \hat{\rho}$, $x(\rho)$ will stay in \mathcal{X}_a and $P^d(\rho) \leq P^d(\hat{\rho})$.

Theorem 2.3. *The canonical dual problem (P^d) is perfectly dual to the primal problem (P) in the sense that if $\hat{\rho} \in \mathcal{S}$ is a KKT point of $P^d(\rho)$, then the vector $\hat{x} = x(\hat{\rho})$ is a KKT point of (P) and*

$$P(\hat{x}) = P^d(\hat{\rho}). \quad (2.14)$$

Proof. By introducing the Lagrange multiplier vector $\lambda \in \mathcal{R}^n$ to relax the inequality constraint $\rho \geq 0$ in \mathcal{S} , the Lagrangian function associated with (P^d) becomes $L(\rho, \lambda) = P^d(\rho) - \lambda^T \rho$. Then the KKT conditions of (P^d) become

$$\begin{aligned} \nabla L &= \nabla P^d(\hat{\rho}) - \lambda = 0, \\ \lambda^T \hat{\rho} &= 0, \quad \lambda \leq 0, \hat{\rho} \geq 0. \end{aligned} \quad (2.15)$$

Notice that $\nabla P^d(\rho) = ((1/2)[x_1^2(\rho) - \ell_1], (1/2)[x_2^2(\rho) - \ell_2], \dots, (1/2)[x_n^2(\rho) - \ell_n])^T$. It follows from conditions (2.15) that $x(\hat{\rho})$ satisfies the complementary conditions of (P) . By the definition of the flow $x(\rho)$, the equation $\nabla P(x(\hat{\rho})) + \text{Diag}(\hat{\rho})x(\hat{\rho}) = 0$ holds. This proved that if $\hat{\rho} \in \mathcal{S}$ is a KKT point of $P^d(\rho)$, then the vector $\hat{x} = x(\hat{\rho})$ defined by (2.6) is a KKT point of the primal problem (P) .

In addition, we have

$$P^d(\hat{\rho}) = P(x(\hat{\rho})) + \sum_{i=1}^n \frac{\hat{\rho}_i}{2} [x_i^2(\hat{\rho}) - \ell_i] = P(x(\hat{\rho})) = P(\hat{x}). \quad (2.16)$$

This completes the proof. \square

Remark 2.4. Theorem 2.3 shows that by using the canonical dual function (2.7), there is no duality gap between the primal problem (\mathcal{P}) and its canonical dual (\mathcal{P}^d), that is, $\theta = 0$. It eliminates this duality gap inherent in the classical Lagrange duality theory and provides necessary conditions for searching global solutions. Actually, we replace \mathcal{S} in (\mathcal{P}^d) with the space $\mathcal{S}_\#, \mathcal{S}_\# := \{\rho \geq 0 \mid \det[\nabla^2 P(x) + \text{Diag}(\rho)] \neq 0, \text{ for all } x \in \mathcal{X}_a\}$ in the proof of Theorem 2.3. Moreover, the inequality of $\det[\nabla^2 P(x) + \text{Diag}(\rho)] \neq 0, \text{ for all } x \in \mathcal{X}_a$ in $\mathcal{S}_\#$ is essentially not a constraint as indicated in [5].

Due to introducing a differential flow $x(\rho)$, the constrained nonconvex problem can be converted to the canonical (perfect) dual problem, which can be solved by deterministic methods. In view of the process (2.2)–(2.6), the flow $x(\rho)$ is based on the KKT (2.2). In other words, we can solve equation (2.2) backwards from ρ^* to get the backward flow $x(\rho), \rho \in \mathcal{S} \cap \{0 \leq \rho \leq \rho^*\}$. Then, it is of interest to know whether there exists a pair (x^*, ρ^*) satisfying (2.2).

Lemma 2.5. *Suppose that $\nabla P(0) \neq 0$. For the primal problem (\mathcal{P}), there exist a point $\rho^* \in \mathcal{S}$ and a nonzero vector $x^* \in \mathcal{X}_a$ such that $\nabla P(x^*) + \text{Diag}(\rho^*)x^* = 0$.*

Proof. Since \mathcal{X}_a is bounded and $P(x)$ is twice continuously differentiable in \mathcal{R}^n , we can choose a large positive real $M \in \mathcal{R}$ such that $\nabla^2 P(x) + \text{Diag}(Me) > 0$, for all $x \in \mathcal{X}_a$ and $\sup_{\mathcal{X}_a} |(\nabla P(x))_i| < M\sqrt{\ell_i}, i = 1, 2, \dots, n$ ($e \in \mathcal{R}^n$ is an n -vector of all ones). Then, it is easy to verify that $\nabla P(x) + \text{Diag}(Me)x < 0$ at the point $x = -\ell^{1/2}$, and $\nabla P(x) + \text{Diag}(Me)x > 0$ at the point $x = \ell^{1/2}$.

Notice that the function $\nabla P(x) + \text{Diag}(Me)x$ is continuous and differentiable in \mathcal{R}^n . It follows from differential and integral calculus that there is a nonzero stationary point $x^* \in \mathcal{X}_a$ such that $\nabla P(x^*) + \text{Diag}(Me)x^* = 0$. Let $\rho^* = Me \in \mathcal{S}$. Thus, there exist a point $\rho^* \in \mathcal{S}$ and a nonzero vector $x^* \in \mathcal{X}_a$ satisfying (2.2). This completes the proof. \square

Remark 2.6. Actually, Lemma 2.5 gives us some information to search the desired parameter ρ^* . From Lemma 2.5, we only need to choose a large positive real $M \in \mathcal{R}$ such that $\nabla^2 P(x) + \text{Diag}(Me) > 0$, for all $x \in \mathcal{X}_a$ and $\sup_{\mathcal{X}_a} |(\nabla P(x))_i| < M\sqrt{\ell_i}, i = 1, 2, \dots, n$. Since $\nabla P(0) \neq 0$, then it follows from $(\|\nabla^2 P(x)\|/M) < 1$ uniformly in \mathcal{X}_a that there is a unique nonzero fixed point $x^* \in \mathcal{X}_a$ such that

$$\frac{-\nabla P(x^*)}{M} = x^* \quad (2.17)$$

which is equivalent to $\nabla P(x^*) + \text{Diag}(Me)x^* = 0$ by Brown fixed-point theorem. In [11, 12], some good algorithms are given to estimate the bounds of $\|\nabla^2 P(x)\|$. If there is a positive real number K such that $\|\nabla^2 P(x)\| \leq K$, for all $x \in \mathcal{X}_a$, then a properly large parameter M can be obtained by the inequalities

$$\frac{\|\nabla^2 P(x)\|}{M} \leq \frac{K}{M} < 1, \quad M > \sup_{\mathcal{X}_a} \left| \frac{(\nabla P(x))_i}{\sqrt{\ell_i}} \right|, \forall i \quad (2.18)$$

uniformly on \mathcal{X}_a for us to use Brown fixed-point theorem. We should choose

$$M > \max \left\{ \sup_{\mathcal{X}_a} \|\nabla^2 P(x)\|, \sup_{\mathcal{X}_a} \left| \frac{(\nabla P(x))_i}{\sqrt{\ell_i}} \right|, \forall i \right\}. \quad (2.19)$$

Finally, let $\rho^* = Me$ which is the desired parameter ρ^* for us. We will discuss to calculate the parameter ρ^* in detail by the use of the results in [15, 16] with the future works.

Remark 2.7. Moreover, for the proper parameter ρ^* , it is worth investigating how to get the solution x^* of (2.2) inside of \mathcal{X}_a . For this issue, when $P(x)$ is a polynomial, we may be referred to [17]. There are results in [17] on bounding the zeros of a polynomial. We may consider for a given bounds to determine the parameter by the use of the results in [17] on the relation between the zeros and the coefficients. We will discuss it with the future works as well. However, the KKT conditions are only necessary conditions for local minimizers to satisfy for the nonconvex case of (\mathcal{P}) . To identify, a global minimizer among all KKT points remains a key task for us to address in the next section.

3. Complete Solutions to Global Optimization Problems

Theorem 3.1. *Suppose that $\hat{\rho}$ is a KKT point of the canonical dual function $P^d(\rho)$ and $\hat{x} = x(\hat{\rho})$ defined by (2.6). If $\hat{\rho} \in \mathcal{S}$, then $\hat{\rho}$ is a global maximizer of (P^d) on \mathcal{S} , and \hat{x} is a global minimizer of (\mathcal{P}) on \mathcal{X}_a and*

$$P(\hat{x}) = \min_{x \in \mathcal{X}_a} P(x) = \max_{\rho \in \mathcal{S}} P^d(\rho) = P^d(\hat{\rho}). \quad (3.1)$$

Proof. If $\hat{\rho} \in \mathcal{S}$ is a KKT point of (P^d) on \mathcal{S} , by (2.15), $x(\hat{\rho})$ stays in \mathcal{X}_a , that is, $(1/2)(x_i^2(\hat{\rho}) - \ell_i) \leq 0$ for all i . By Lemma 2.1 and 2.2, it is easy to verify that $x(\rho) \in \mathcal{X}_a$ and $P^d(\hat{\rho}) \geq P^d(\rho)$ for any $\rho \geq \hat{\rho}$.

For any given parameter ρ , ($\rho \geq \hat{\rho}$), we define the function $f_\rho(x)$ as follows:

$$f_\rho(x) = P(x) + \sum_{i=1}^n \frac{\rho_i}{2} (x_i^2 - \ell_i). \quad (3.2)$$

It is obvious that $P(x) \geq f_\rho(x)$ for all $x \in \mathcal{X}_a$. Since $f_\rho(x)$ is twice continuously differentiable in \mathcal{R}^n , there exists a closed convex region \mathcal{E} containing \mathcal{X}_a such that on \mathcal{E} ,

$$\begin{aligned} \nabla \left\{ P(x(\rho)) + \sum_{i=1}^n \frac{\rho_i}{2} [x_i^2(\rho) - \ell_i] \right\} &= \nabla P(x(\rho)) + \text{Diag}(\rho)x(\rho) = 0, \\ \nabla^2 \left\{ P(x) + \sum_{i=1}^n \frac{\rho_i}{2} (x_i^2 - \ell_i) \right\} &= \nabla^2 P(x) + \text{Diag}(\rho) > 0, \quad \forall x \in \mathcal{X}_a. \end{aligned} \quad (3.3)$$

This implies that $x(\rho)$ is the unique global minimizer of $f_\rho(x)$ over \mathcal{X} . By (2.7), we have

$$f_\rho(x(\rho)) = P(x(\rho)) + \sum_{i=1}^n \frac{\rho_i}{2} [x_i^2(\rho) - \ell_i] = P^d(\rho). \quad (3.4)$$

Thus, for any $\rho \geq \hat{\rho}$,

$$P(x) \geq f_\rho(x) \geq \min_{x \in \mathcal{X}_a} f_\rho(x) = f_\rho(x(\rho)) = P^d(\rho). \quad (3.5)$$

On the other hand, by the fact that the canonical dual function $P^d(\rho)$ is concave on \mathcal{S} , $\hat{\rho}$ must be a global maximizer of (P^d) on \mathcal{S} , and we have

$$\max_{\rho \in \mathcal{S}} P^d(\rho) = P^d(\hat{\rho}) = P(x(\hat{\rho})) + \sum_{i=1}^n \frac{\hat{\rho}_i}{2} [x_i^2(\hat{\rho}) - \ell_i] = P(x(\hat{\rho})), \quad (3.6)$$

and for all $x \in \mathcal{X}_a$,

$$P(x) \geq \max_{\rho \geq \hat{\rho}} P^d(\rho) = P^d(\hat{\rho}) = P(\hat{x}). \quad (3.7)$$

Thus, \hat{x} is a global minimizer of (\mathcal{D}) on \mathcal{X}_a and

$$P(\hat{x}) = \min_{x \in \mathcal{X}_a} P(x) = \max_{\rho \in \mathcal{S}} P^d(\rho) = P^d(\hat{\rho}). \quad (3.8)$$

This completes the proof. \square

Remark 3.2. Theorem 3.1 shows that a vector $\hat{x} = x(\hat{\rho})$ is a global minimizer of the problem (\mathcal{D}) if $\hat{\rho} \in \mathcal{S}$ is a critical solution of (P^d) . However, for certain given $P(x)$, the canonical dual function might have no critical point in \mathcal{S} . For example, the canonical dual solutions could locate on the boundary of \mathcal{S} . In this case, the primal problem (\mathcal{D}) may have multiple solutions.

In order to study the existence conditions of the canonical dual solutions, we let $\partial\mathcal{S}$ denote the boundary of \mathcal{S} .

Theorem 3.3. *Suppose that $P(x)$ is a given twice continuously differentiable function, $\mathcal{S} \neq \emptyset$ and $\partial\mathcal{S} \neq \emptyset$. If for any given $\rho_0 \in \partial\mathcal{S}$ and $\rho \in \mathcal{S}$,*

$$\lim_{\alpha \rightarrow 0^+} P^d(\rho_0 + \alpha\rho) = -\infty, \quad (3.9)$$

then the canonical dual problem (P^d) has a critical point $\hat{\rho} \in \mathcal{S}$, and $x(\hat{\rho})$ is a global minimizer to (\mathcal{D}) .

Proof. We first show that for any given $\bar{\rho} \geq 0 \in \mathcal{R}^n$, $\bar{\rho} \neq 0$,

$$\lim_{\alpha \rightarrow \infty} P^d(\alpha\bar{\rho}) = -\infty. \quad (3.10)$$

Notice that there exist a point $\rho^* \in \mathcal{S}$ and a nonzero vector $x(\rho^*) \in \mathcal{X}_a$. For any given $\bar{\rho}$, the inequality $\alpha_0 \bar{\rho} > \rho^*$ always holds as $\alpha_0 > 0$ becomes large enough. Then for any $\alpha \geq \alpha_0$, it follows from Lemma 2.1 and 2.2 that $\alpha \bar{\rho} \in \mathcal{S}$ and $x(\alpha \bar{\rho})$ stays in \mathcal{X}_a , that is, $x_i^2(\alpha \bar{\rho}) \leq \ell_i$ for all i . It means that there exists a large positive real L such that $|P(\alpha \bar{\rho})| \leq L$ for $\alpha \geq \alpha_0$ since $P(x)$ is twice continuously differentiable in \mathcal{R}^n .

By Lemma 2.2, we have

$$\begin{aligned} \frac{dP^d(\alpha \bar{\rho})}{d\alpha} &= \left(\nabla P^d(\alpha \bar{\rho}) \right)^T \bar{\rho} = \sum_{i=1}^n \frac{\bar{\rho}_i}{2} \left[x_i^2(\alpha \bar{\rho}) - \ell_i \right], \\ \frac{d^2 P^d(\alpha \bar{\rho})}{d\alpha^2} &= -\bar{\rho}^T \text{Diag}(x(\alpha \bar{\rho})) \left[\nabla^2 P(x(\alpha \bar{\rho})) + \text{Diag}(\alpha \bar{\rho}) \right]^{-1} \text{Diag}(x(\alpha \bar{\rho})) \bar{\rho}, \end{aligned} \quad (3.11)$$

where $\text{Diag}(x(\alpha \bar{\rho})) \bar{\rho} = (\bar{\rho}_1 x_1(\alpha \bar{\rho}), \bar{\rho}_2 x_2(\alpha \bar{\rho}), \dots, \bar{\rho}_n x_n(\alpha \bar{\rho}))^T$. For any $\alpha \geq \alpha_0$, by the definition of \mathcal{S} , it is easy to see that $d^2 P^d(\alpha \bar{\rho})/d\alpha^2 \leq 0$, namely, $dP^d(\alpha \bar{\rho})/d\alpha$ monotonously decreases on $[\alpha_0, +\infty)$. Moreover, since $x(\alpha \bar{\rho}) \in \mathcal{X}_a$, we have $dP^d(\alpha \bar{\rho})/d\alpha \leq 0$ by (3.11). Then, $P^d(\alpha \bar{\rho})$ is monotonously decreasing on $[\alpha_0, +\infty)$. Thus, to prove (3.10), it is only needed to show that there exists a positive real $\bar{\alpha} \geq \alpha_0$ such that

$$\frac{dP^d(\bar{\alpha} \bar{\rho})}{d\bar{\alpha}} = \sum_{i=1}^n \frac{\bar{\rho}_i}{2} \left[x_i^2(\bar{\alpha} \bar{\rho}) - \ell_i \right] < 0, \quad (3.12)$$

which implies that $P^d(\bar{\alpha} \bar{\rho})$ is strictly monotonously decreasing on $[\bar{\alpha}, +\infty)$.

Suppose that

$$\sum_{i=1}^n \frac{\bar{\rho}_i}{2} \left[x_i^2(\alpha_1 \bar{\rho}) - \ell_i \right] = 0 \quad (3.13)$$

at a point $\alpha_1 \geq \alpha_0$. Since $\bar{\rho}_i [x_i^2(\alpha_1 \bar{\rho}) - \ell_i] \leq 0$ for all i and $\bar{\rho} \neq 0$, the equation $x_r^2(\alpha_1 \bar{\rho}) = \ell_r$ holds for some subscript r , which means that $\bar{\rho}_r x_r(\alpha_1 \bar{\rho}) \neq 0$. By positive definiteness of $[\nabla^2 P(x(\alpha_1 \bar{\rho})) + \text{Diag}(\alpha_1 \bar{\rho})]$, we have $dP^d(\alpha_1 \bar{\rho})/d\alpha_1 = 0$ and $d^2 P^d(\alpha_1 \bar{\rho})/d\alpha_1^2 < 0$. One can verify that $dP^d(\alpha \bar{\rho})/d\alpha < 0$ on $(\alpha_1, +\infty)$ since its monotonicity. Obviously, it is very easy to hold (3.12) if there does not exist α satisfying (3.13). Consequently, there always exists a positive real $\bar{\alpha}$ such that $P^d(\bar{\alpha} \bar{\rho})$ is strictly monotonously decreasing on $[\bar{\alpha}, +\infty)$. It leads to the conclusion (3.10).

Since $P^d : \mathcal{S} \rightarrow \mathcal{R}$ is concave and the condition (3.10) holds, if (3.9) holds, then the canonical dual function $P^d(\rho)$ is coercive on the open convex set \mathcal{S} . Therefore, the canonical dual problem (P^d) has one maximizer $\hat{\rho} \in \mathcal{S}$ by the theory of convex analysis [4, 18]. This completes the proof. \square

Clearly, when $\nabla^2 P(x) > 0$ on \mathcal{X}_a , the dual feasible space \mathcal{S} is equivalent to $\mathcal{R}_+^n = \{\rho \in \mathcal{R}^n \mid \rho \geq 0\}$ and $\text{Diag}(\rho) \in \mathcal{R}^{n \times n}$ by (2.1). Notice that $\lim_{\alpha \rightarrow \infty} P^d(\alpha \bar{\rho}) = -\infty$ for any given $\bar{\rho} \geq 0$. Then $P^d(\rho)$ is concave and coercive on \mathcal{R}_+^n and (P^d) has at least one maximizer on \mathcal{R}_+^n . In this case, it is then of interest to characterize a unique solution of (\mathcal{D}) by the dual variable.

Let

$$I = \left\{ i \in \{1, 2, \dots, n\} \mid x_i^2(0) < \ell_i \right\}, \quad J = \frac{\{1, 2, \dots, n\}}{I}. \quad (3.14)$$

Theorem 3.4. *If $\nabla^2 P(x) > 0$ on \mathcal{X}_a , the primal problem (\mathcal{P}) has a unique global minimizer $x(\hat{\rho})$ determined by $\hat{\rho} \in \mathcal{R}_+^n$ satisfying*

$$\hat{\rho}_i = 0, \quad \forall i \in I, \quad x_i^2(\hat{\rho}) = \ell_i, \quad \forall i \in J. \quad (3.15)$$

Proof. To prove Theorem 3.4, by Theorem 2.3 and 3.1, it is only needed to prove that $\hat{\rho}$ is a KKT point of (P^d) in \mathcal{R}_+^n . By (3.14), the relations $x_i^2(\hat{\rho}) < \ell_i$ for all $i \in I$ also hold. Since $\hat{\rho}$ satisfies equations $x_i^2(\hat{\rho}) = \ell_i$ for all $i \in J$, we can verify that $x(\hat{\rho})$ stays in \mathcal{X}_a and the complementarity conditions $\hat{\rho}_i(x_i^2(\hat{\rho}) - \ell_i) = 0$, for all i hold. Thus, $\hat{\rho}$ is a KKT point of (P^d) in \mathcal{R}_+^n by (2.9), (2.15), and $x(\hat{\rho})$ is a unique global minimizer of (\mathcal{P}) . This completes the proof. \square

Before beginning of applications to optimal control problems, we present two examples to find global minimizers by differential flows.

Example 3.5. As a particular example of (\mathcal{P}) , let us consider the following one dimensional nonconvex minimization problem with a box:

$$\begin{aligned} \min \quad & P(x) = \frac{1}{3}x^3 + 2x \\ \text{s.t.} \quad & x^2 \leq 1. \end{aligned} \quad (3.16)$$

We have $P'(x) = x^2 + 2$ and $P''(x) = 2x$, for all $x^2 \leq 1$. By choosing $\rho^* = 6\sqrt{2}$, we solve the following equation in $\{x^2 \leq 1\}$:

$$x^2 + 2 + 6\sqrt{2}x = 0 \quad (3.17)$$

to get a solution $x^* = -2/(4 + 3\sqrt{2})$. Next we solve the following boundary value problem of the ordinary differential equation:

$$\dot{x} = \frac{-x}{2x + \rho}, \quad x(6\sqrt{2}) = \frac{-2}{4 + 3\sqrt{2}}, \quad 2 < \rho \leq 6\sqrt{2}. \quad (3.18)$$

To find a parameter such that

$$x^2(\rho) = 1, \quad (3.19)$$

we get

$$\hat{\rho} = 3, \quad (3.20)$$

which satisfies

$$P''(x) + \hat{\rho} = 2x + 3 > 0, \quad \forall x^2 \leq 1. \quad (3.21)$$

Let $x(3)$ be denoted by \hat{x} . To find the value of \hat{x} , we compute the solution of the following algebra equation:

$$x^2 + 2 + 3x = 0, \quad x^2 = 1 \quad (3.22)$$

and get $\hat{x} = -1$. It follows from Theorem 3.1 that $\hat{x} = -1$ is a global minimizer of Example 3.5.

Remark 3.6. In this example, we see that a differential flow is useful in solving a nonconvex optimization problem. For the global optimization problem, people usually compute the global minimizer numerically. Even in using canonical duality method, one has to solve a canonical dual problem numerically. Nevertheless, the differential flow directs us to a new way for finding a global minimizer. Particularly, one may expect an exact solution of the problem provided that the corresponding differential equation has an analytic solution.

Example 3.7. Given a symmetric matrix $A \in \mathcal{R}^{n \times n}$ and a vector $c \in \mathcal{R}^n$. Let $P(x) = (1/2)x^T Ax - c^T x$. We consider the following box-constrained nonconvex global optimization:

$$\begin{aligned} \min \quad & P(x) = \frac{1}{2}x^T Ax - c^T x \\ \text{s.t.} \quad & -\sqrt{\ell_i} \leq x_i \leq \sqrt{\ell_i}, \quad i = 1, 2, \dots, n. \end{aligned} \quad (3.23)$$

Since A is an indefinite matrix, we choose a large $\rho^* \in \mathcal{R}^n$ such that $A + \text{Diag}(\rho^*) > 0$ and $\sup_{-\sqrt{\ell_i} \leq x_i \leq \sqrt{\ell_i}} |(Ax - c)_i| < \rho_i^* \sqrt{\ell_i}$, $i = 1, 2, \dots, n$. We see that the differential equation is

$$\begin{aligned} dx_i(\rho) &= [A + \text{Diag}(\rho)]_{i1}(\rho)x_1(\rho)d\rho_1 + [A + \text{Diag}(\rho)]_{i2}(\rho)x_2(\rho)d\rho_2 \\ &\quad + \dots + [A + \text{Diag}(\rho)]_{in}(\rho)x_n(\rho)d\rho_n, \\ x(\rho^*) &= [A + \text{Diag}(\rho^*)]^{-1}c, \quad \rho \in \mathcal{S} \cap \{0 \leq \rho \leq \rho^*\}, \end{aligned} \quad (3.24)$$

where $\mathcal{S} = \{\rho \geq 0 \mid A + \text{Diag}(\rho) > 0\}$. It leads a differential flow

$$x(\rho) = [A + \text{Diag}(\rho)]^{-1}c, \quad \rho \in \mathcal{S}. \quad (3.25)$$

For simplicity without generality, we assume that $A = \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix}$ and $c = \begin{pmatrix} c_1 \\ c_1 \end{pmatrix}$. If for ρ , $\det[A + \text{Diag}(\rho)] \neq 0$, we have

$$[A + \text{Diag}(\rho)]^{-1} = \begin{pmatrix} \frac{a_3 + \rho_2}{(a_1 + \rho_1)(a_3 + \rho_2) - a_2^2} & \frac{-a_2}{(a_1 + \rho_1)(a_3 + \rho_2) - a_2^2} \\ \frac{-a_2}{(a_1 + \rho_1)(a_3 + \rho_2) - a_2^2} & \frac{a_1 + \rho_1}{(a_1 + \rho_1)(a_3 + \rho_2) - a_2^2} \end{pmatrix}. \quad (3.26)$$

Then the dual problem can be formulated as

$$\max \left\{ P(\rho) = -\frac{1}{2}c^T [A + \text{Diag}(\rho)]^{-1}c - \ell^T \rho : \rho \in \mathcal{S} \right\}. \quad (3.27)$$

If we choose $a_1 = -0.5, a_2 = -0.5, a_3 = -0.3, c_1 = c_2 = 0.3$ and $\ell_1 = 0.5, \ell_2 = 2$, this dual problem has only one KKT point $\hat{\rho} = (1.8, 0.7)^T \in \mathcal{S}$. By Theorem 3.1, $\hat{x} = [A + \text{Diag}(\hat{\rho})]^{-1}c = (1, 2)^T$ is a global minimizer of Example 3.7 and $P(1, 2) = -2.75 = P^d(1.8, 0.7)$.

4. Applications to Constrained Optimal Control Problems

In this section, we consider the following constrained linear-quadratic optimal control problem:

$$\begin{aligned} \min \quad & J(x, u) = \frac{1}{2} \int_0^T x^T(t)Qx(t) + u^T(t)Ru(t)dt \\ \text{s.t.} \quad & \dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, u(t) \in U, t \in [0, T], \end{aligned} \quad (4.1)$$

where $Q \in \mathcal{R}^{n \times n}, R \in \mathcal{R}^{m \times m}$ are positive semidefinite and positive definite symmetric matrices respectively, $x(t) \in \mathcal{R}^n$ is a state vector, and $u(t) \in \mathcal{R}^m$ is integrable or piecewise continuous on $[0, T]$ within U . Simply, $U = \{u \in \mathcal{R}^m \mid -1 \leq u_i \leq 1\}$, and U is a unit box. Problems of the above type arise naturally in system science and engineering with wide applications [19, 20].

It is well known that the central result in the optimal control theory is the *Pontryagin maximum principle* providing necessary conditions for optimality in very general optimal control problems.

4.1. Pontryagin Maximum Principle

Define the Hamilton-Jacobi-Bellman function

$$H(x, u, \lambda) = \lambda^T (Ax + Bu) + \frac{1}{2}x^T Qx + \frac{1}{2}u^T Ru. \quad (4.2)$$

If the control \hat{u} is an optimal solution for the problem (4.1), with $\hat{x}(\cdot)$ and $\hat{\lambda}(\cdot)$ denoting the state and costate corresponding to $\hat{u}(\cdot)$, respectively, then \hat{u} is an extremal control, that is, we have

$$\begin{aligned} \dot{\hat{x}} &= H_\lambda(\hat{x}, \hat{u}, \hat{\lambda}) = A\hat{x} + B\hat{u}, & \hat{x}(0) &= x_0, \\ \dot{\hat{\lambda}} &= -H_x(\hat{x}, \hat{u}, \hat{\lambda}) = -A^T\hat{\lambda} - Q\hat{x}, & \hat{\lambda}(T) &= 0, \end{aligned} \quad (4.3)$$

and a.e. $t \in [0, T]$,

$$H(t, \hat{x}(t), \hat{u}(t), \hat{\lambda}(t)) = \min_{u \in U} H(t, \hat{x}(t), u(t), \hat{\lambda}(t)). \quad (4.4)$$

Unfortunately, above conditions are not, in general, sufficient for optimality. In such a case, we need to go through the process of comparing all the candidates for optimality that the necessary conditions produce, and picking out an optimal solution to the problem. Nevertheless, Lemma 4.1 can prove that the solution satisfies sufficiency conditions of the type considered in this section, then these conditions will ensure the optimality of the solution.

Lemma 4.1. *Let $\hat{u}(\cdot)$ be an admissible control, $\hat{x}(\cdot)$ and $\hat{\lambda}(\cdot)$ be the corresponding state and costate. If $\hat{x}(t)$, $\hat{u}(t)$, and $\hat{\lambda}(t)$ satisfy the Pontryagin maximum principle ((4.3)-(4.4)), then $\hat{u}(t)$ is an optimal control to the problem (4.1).*

Proof. For any given x, λ , let

$$H^*(x, \lambda) = \min_{u \in U} H(x, u, \lambda). \quad (4.5)$$

For any $u \in U$, by the definition of H^* , $H^*(x, \lambda) \leq H(x, u, \lambda)$, and $H^*(x, \lambda)$ is equivalent to the following global optimization

$$\min_{u \in U} \frac{1}{2} u^T R u + \lambda^T B u. \quad (4.6)$$

Moreover, we can derive an analytic form of the global minimizer for (4.6) via the co-state λ . It is easy to see that the minimizer \hat{u} of (4.6) doesn't depend on x , that is, $\partial \hat{u} / \partial x = 0$ which implies that

$$H_x^*(x, \lambda) = H_x(x, \hat{u}, \lambda) + H_u(x, \hat{u}, \lambda) \frac{\partial \hat{u}}{\partial x} = H_x(x, \hat{u}, \lambda). \quad (4.7)$$

Since U is a closed convex set, by the classical linear systems theory, the state set X of (4.1) is a convex subset of \mathcal{R}^n . By the fact that the minimizer \hat{u} does not depend on x and the convexity of the integrand in the cost functional, the function $H^*(x, \lambda)$ is convex with respect to x over X . In other words, for any $x \in X$, and a.e. $t \in [0, T]$,

$$\begin{aligned} H^*(x, \hat{\lambda}(t)) &\geq H^*(\hat{x}(t), \hat{\lambda}(t)) + H_x^{*T}(\hat{x}(t), \hat{\lambda}(t))(x - \hat{x}(t)) \\ &= H(\hat{x}(t), \hat{u}(t), \hat{\lambda}(t)) + H_x^T(\hat{x}(t), \hat{u}(t), \hat{\lambda}(t))(x - \hat{x}(t)). \end{aligned} \quad (4.8)$$

Thus, for any admissible pair $(x(\cdot), u(\cdot))$, and a.e. $t \in [0, T]$, by (4.5), we have

$$\begin{aligned} H^*(x(t), u(t), \hat{\lambda}(t)) &\geq H^*(x(t), \hat{\lambda}(t)) \\ &\geq H(\hat{x}(t), \hat{u}(t), \hat{\lambda}(t)) + H_x^T(\hat{x}(t), \hat{u}(t), \hat{\lambda}(t))(x(t) - \hat{x}(t)), \end{aligned} \quad (4.9)$$

which leads to

$$H(\hat{x}(t), \hat{u}(t), \hat{\lambda}(t)) - H(x(t), u(t), \hat{\lambda}(t)) \leq \hat{\lambda}^T(t)(x(t) - \hat{x}(t)). \quad (4.10)$$

Notice that $\hat{\lambda}(T) = 0$ and $x(0) = \hat{x}(0) = x_0$. We can obtain

$$\begin{aligned} J[\hat{u}(\cdot)] - J[u(\cdot)] &= \int_0^T \left[H(\hat{x}(t), \hat{u}(t), \hat{\lambda}(t)) - H(x(t), u(t), \hat{\lambda}(t)) \right] dt \\ &\quad + \int_0^T \hat{\lambda}^T(t) (\dot{x}(t) - \hat{\dot{x}}(t)) dt \leq \int_0^T d[\hat{\lambda}^T(t)(x(t) - \hat{x}(t))] \\ &= -\hat{\lambda}^T(0)(x(0) - \hat{x}(0)) = 0. \end{aligned} \quad (4.11)$$

This means that J attains its minimum at \hat{u} . The proof is completed. \square

Lemma 4.1 reformulates the constrained optimal control problem (4.1) into a global optimization problem (4.6). Based on Theorem 3.4, an analytic solution of (4.1) can be expressed via the co-state.

Theorem 4.2. *Suppose that*

$$u = \arg \min_{u \in U} \frac{1}{2} u^T R u + \lambda^T B u. \quad (4.12)$$

We have the following expression

$$u = -[R + \text{Diag}(\rho(\lambda))]^{-1} B^T \lambda, \quad (4.13)$$

where $\rho(\lambda)$ with respect to the co-state λ is given by $\rho(\lambda) \geq 0$ satisfying

$$\begin{aligned} \rho_i(\lambda) &= 0 \quad \text{if } (R^{-1} B^T \lambda)_i^2 < 1, \\ \left[(R + \text{Diag}(\rho(\lambda)))^{-1} B^T \lambda \right]_i^2 &= 1 \quad \text{if } (R^{-1} B^T \lambda)_i^2 \geq 1. \end{aligned} \quad (4.14)$$

Proof. The proof of Theorem 4.2 is parallel to Theorem 3.4. \square

Substituting $u = -[R + \text{Diag}(\rho(\lambda))]^{-1} B^T \lambda$ into (4.3), we have

$$\begin{aligned} \dot{x} &= Ax + B \left(-[R + \text{Diag}(\rho(\lambda))]^{-1} B^T \lambda \right), \quad x(0) = x_0, \\ \dot{\lambda} &= -A^T \lambda - Qx, \quad \lambda(T) = 0. \end{aligned} \quad (4.15)$$

If $(\hat{x}(\cdot), \hat{\lambda}(\cdot))$ is a solution of the above equations (4.15), let

$$\hat{u}(\cdot) = -\left[R + \text{Diag}(\rho(\hat{\lambda}(\cdot))) \right]^{-1} B^T \hat{\lambda}(\cdot). \quad (4.16)$$

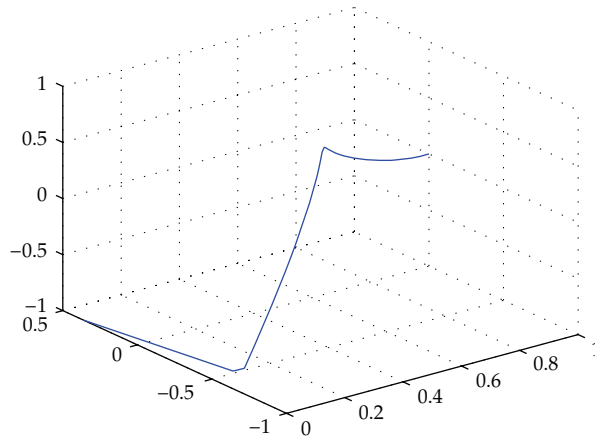


Figure 1: The optimal control \hat{u} in Example 4.3.

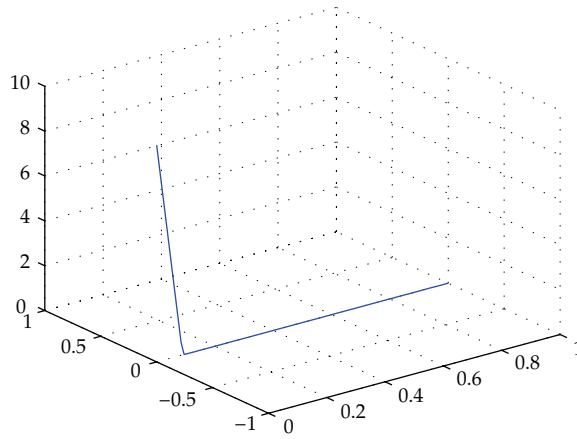


Figure 2: The dual variable $\rho(\hat{\lambda})$ in Example 4.3.

By Lemma 4.1, $\hat{x}(\cdot), \hat{u}(\cdot), \hat{\lambda}(\cdot)$ satisfy the Pontryagin maximum principle, and we present an analytic form of the optimal control to (4.1) via the canonical dual variable

$$\hat{u}(t) = -\left[R + \text{Diag}\left(\rho\left(\hat{\lambda}(t)\right)\right)\right]^{-1} B^T \hat{\lambda}(t), \quad \text{a.e. } t \in [0, T]. \tag{4.17}$$

Next, we give an example to illustrate our results.

Example 4.3. We consider

$$A = \begin{pmatrix} 2 & 7 \\ -2.5 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -5 \\ 1.5 & 7 \end{pmatrix}, \quad Q = \begin{pmatrix} 2.4 & 3 \\ -2 & -1 \end{pmatrix}, \quad R = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad x(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{4.18}$$

and $T = 1$ in (4.1). $Q \geq 0$ and $R > 0$ satisfy the assumption in (4.1).

Following the idea of Lemma 4.1 and Theorem 4.2, we need to solve the following boundary value problem for differential equations to derive the optimal solution

$$\begin{aligned}\dot{\hat{x}} &= H_\lambda(\hat{x}, \hat{u}, \hat{\lambda}) = A\hat{x} + B\hat{u}, \quad \hat{x}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ \dot{\hat{\lambda}} &= -H_x(\hat{x}, \hat{u}, \hat{\lambda}) = -A^T\hat{\lambda} - Q\hat{x}, \quad \hat{\lambda}(1) = 0, \\ \hat{u}_1 &= -\frac{(B^T\hat{\lambda})_1}{\max[2, |(B^T\hat{\lambda})_1|]}, \quad \hat{u}_2 = -\frac{(B^T\hat{\lambda})_2}{\max[3, |(B^T\hat{\lambda})_2|]}, \quad \text{a.e. } t \in [0, 1].\end{aligned}\tag{4.19}$$

By solving equations (4.19) in MATLAB, we can obtain the optimal control \hat{u} and the dual variable $\rho(\hat{\lambda})$ as follows (see Figures 1 and 2).

Acknowledgments

The authors would like to thank the referees for their helpful comments on the early version of this paper. This work was supported by the National Natural Science Foundation of China (no. 10971053).

References

- [1] C. A. Floudas, *Deterministic Global Optimization: Theory, Methods and Applications*, vol. 37 of *Nonconvex Optimization and Its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
- [2] D. Y. Gao, "Nonconvex semi-linear problems and canonical duality solutions," in *Advances in Mechanics and Mathematics. Vol. II*, vol. 4 of *Adv. Mech. Math.*, pp. 261–312, Kluwer Academic Publishers, Boston, Mass, USA, 2003.
- [3] S.-C. Fang, D. Y. Gao, R.-L. Sheu, and S.-Y. Wu, "Canonical dual approach to solving 0-1 quadratic programming problems," *Journal of Industrial and Management Optimization*, vol. 4, no. 1, pp. 125–142, 2008.
- [4] R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, NJ, USA, 1970.
- [5] D. Y. Gao, "Solutions and optimality criteria to box constrained nonconvex minimization problems," *Journal of Industrial and Management Optimization*, vol. 3, no. 2, pp. 293–304, 2007.
- [6] D. Y. Gao and N. Ruan, "Solutions to quadratic minimization problems with box and integer constraints," *Journal of Global Optimization*, vol. 47, no. 3, pp. 463–484, 2010.
- [7] Z. Wang, S.-C. Fang, D. Y. Gao, and W. Xing, "Global extremal conditions for multi-integer quadratic programming," *Journal of Industrial and Management Optimization*, vol. 4, no. 2, pp. 213–225, 2008.
- [8] D. Y. Gao, "Perfect duality theory and complete solutions to a class of global optimization problems," *Optimization*, vol. 52, no. 4-5, pp. 467–493, 2003.
- [9] D. Y. Gao, *Duality Principles in Nonconvex Systems: Theory, Methods and Applications*, vol. 39 of *Nonconvex Optimization and Its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
- [10] D. Y. Gao, "Canonical dual transformation method and generalized triality theory in nonsmooth global optimization," *Journal of Global Optimization*, vol. 17, no. 1–4, pp. 127–160, 2000.
- [11] C. S. Adjiman, S. Dallwig, C. A. Floudas, and A. Neumaier, "A global optimization method, α BB, for general twice-differentiable constrained NLPs-I. Theoretical advances," *Computers and Chemical Engineering*, vol. 22, no. 9, pp. 1137–1158, 1998.

- [12] C. S. Adjiman, I. P. Androulakis, and C. A. Floudas, "A global optimization method, α BB, for general twice-differentiable constrained NLPs-II. Implementation and computational results," *Computers and Chemical Engineering*, vol. 22, no. 9, pp. 1159–1179, 1998.
- [13] H. Grassmann, *Extension Theory*, vol. 19 of *History of Mathematics*, American Mathematical Society, Providence, RI, USA, 2000.
- [14] G. C. Rota, "Extension theory of differential operators. I," *Communications on Pure and Applied Mathematics*, vol. 11, pp. 23–65, 1958.
- [15] C. S. Adjiman, I. P. Androulakis, and C. A. Floudas, "Global optimization of mixed-integer nonlinear problems," *AIChE Journal*, vol. 46, no. 9, pp. 1769–1797, 2000.
- [16] M. Mönnigmann, "Efficient calculation of bounds on spectra of Hessian matrices," *SIAM Journal on Scientific Computing*, vol. 30, no. 5, pp. 2340–2357, 2008.
- [17] D. Hertz, C. S. Adjiman, and C. A. Floudas, "Two results on bounding the roots of interval polynomials," *Computers and Chemical Engineering*, vol. 23, no. 9, pp. 1333–1339, 1999.
- [18] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, The Netherlands, 1976.
- [19] J. Casti, "The linear-quadratic control problem: some recent results and outstanding problems," *SIAM Review*, vol. 22, no. 4, pp. 459–485, 1980.
- [20] C. Robinson, *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*, Studies in Advanced Mathematics, CRC Press, Boca Raton, Fla, USA, 1995.