Optimizing a Linear Function over a Noncompact Real Algebraic Variety

Feng Guo
School of Mathematical Sciences,
Dalian University of Technology
Dalian, 116024, China
fguo@dlut.edu.cn

Chu Wang
Key Laboratory of Mathematics Mechanization, AMSS
Beijing 100190, China
cwang@mmrc.iss.ac.cn

Lihong Zhi
Key Laboratory of Mathematics Mechanization, AMSS
Beijing 100190, China
lzhi@mmrc.iss.ac.cn

Categories and Subject Descriptors

I.1.2 [**Symbolic and Algebraic Manipulation**]: Algorithms; G.1.6 [**Numerical Analysis**]: Convex programming

General Terms

Algorithms

Keywords

dual variety, noncompact real algebraic variety, pointed convex set, polar

1. INTRODUCTION

Consider the problem of optimizing a linear function over a real algebraic variety $X \subseteq \mathbb{R}^n$

$$\begin{aligned} c_0^* = \max \quad & c_1x_1 + \ldots + c_nx_n \\ \text{s.t.} \quad & x \in X = \{v \in \mathbb{R}^n \mid h_1(v) = \cdots = h_p(v) = 0\}, \end{aligned} \tag{1.1}$$

where $h_1, h_2, \ldots, h_p$ are polynomials in unknowns $x_1, \ldots, x_n$. There exists a polynomial $\Phi(c_0, c_1 \ldots, c_n)$ in $n+1$ variables such that

$$\Phi(c_0^*, c_1, \ldots, c_n) = 0.$$

Our aim is to compute such a polynomial Φ of the least possible degree.

In [3, 4], Rostalski and Sturmfels explored dualities and their interconnections in the context of polynomial optimization (1.1). Assuming that the feasible region X is irreducible, compact and smooth, they showed that the optimal value function Φ is represented by the defining equation of the hypersurface dual to the projective closure of X [4, Theorem 5.23]. In the present paper, we prove this conclusion is still true for a noncompact real algebraic variety X, when X is irreducible, smooth and the recession cone of the closure of the convex hull $\mathbf{co}\,(X)$ of X is pointed.

SNC'14, July 28-31, 2014, Shanghai, China.

http://dx.doi.org/10.1145/2631948.2631957

2. THE NON-COMPACT POINTED CASE

Let C be a non-empty convex set in $\mathbb{R}^n$. We denote $\mathbf{cl}\,(C)$, $\mathbf{int}\,(C)$ and $\mathbf{ri}\,(C)$ as the closure, interior and relative interior of C respectively. The support function of C is defined by

$$\delta^*(x^* \mid C) = \sup\{\langle x, x^*\rangle \mid x \in C\}.$$

Let $\mathbf{dom}\,(\delta^*(x^* \mid C))$ denote the effective domain of $\delta^*(x^* \mid C)$. Note that $\delta^*(x^* \mid C)$ is a proper convex function and $\mathbf{dom}\,(\delta^*(x^* \mid C))$ consists of all vectors $a \in \mathbb{R}^n$ such that the maximal value of $f(x) = a^Tx$ on C is finite.

The *recession cone* 0^+C of C is the set including all vectors y satisfying $x + \lambda y \in C$ for every $\lambda > 0$ and $x \in C$.

Proposition 2.1. *Let C_1 and C_2 be closed convex cone with $C_2 \subseteq C_1$ and $C_2 \cap \mathbf{ri}\,(C_1) \neq \emptyset$, then $0^+(C_2) \subseteq 0^+(C_1)$.*

Proof. It is a corollary of Theorem 8.3 in [2]. □

A convex cone K is *pointed* if it is closed and $K \cap -K = \{0\}$. The *polar* of a non-empty convex cone K is defined as

$$K^\circ = \{x^* \in \mathbb{R}^n \mid \forall x \in K, \langle x, x^*\rangle \leq 0\}.$$

Theorem 2.2. *Let $C \subseteq \mathbb{R}^n$ be a closed and unbounded convex set. Suppose that 0^+C is pointed, then*

(a) $(0^+C)^\circ$ *is an n-dimensional convex set;*

(b) $\mathbf{int}\,\big((0^+C)^\circ\big) \subseteq \mathbf{dom}\,(\delta^*(x^* \mid C)) \subseteq (0^+C)^\circ$. *Moreover, $f(x) = a^Tx$ attains its maximal value on C for every $a \in \mathbf{int}\,\big((0^+C)^\circ\big)$.*

Proof. (a). Since 0^+C is closed by Theorem 8.2 in [2], the conclusion follows from the result in [1, Section 3.3, Exercise 20].

(b). We first show that the maximal value of $f(x) = a^Tx$ on C is finite and attainable for every $a \in \mathbf{int}\,\big((0^+C)^\circ\big)$. Fix a vector $a \in \mathbf{int}\,\big((0^+C)^\circ\big)$, then $f(x) < 0$ for all $x \in 0^+C\backslash\{0\}$.

Suppose the maximal value of $f(x)$ on C is $+\infty$. Let $P = \{x \in \mathbb{R}^n \mid f(x) \geq 0\}$, then $P \cap C$ is unbounded. Moreover, $P \cap \mathbf{ri}\,(C) \neq \emptyset$, otherwise, according to Theorem 6.3 in [2], $\mathbf{cl}\,(C) = \mathbf{cl}\,(\mathbf{ri}\,(C))$, we have $C \subseteq \{x \in \mathbb{R}^n \mid f(x) \leq 0\}$. It contradicts to the assumption that f is not bounded above. Choose a vector $x_0 \in P \cap \mathbf{ri}\,(C)$. Since $P \cap C$ is unbounded, by Theorem 8.4 in [2], there exists a non-zero vector $y \in 0^+(P \cap C)$, $x_0 + \lambda y \in P \cap C$ for all $\lambda \geq 0$, i.e., $f(x_0 + \lambda y) \geq 0$ for all $\lambda \geq 0$. By Proposition 2.1, we have $y \in 0^+C$ and $f(y) < 0$. Therefore $f(x_0 + \lambda y) < 0$ for some $\lambda > 0$, which is a contradiction to the assumption that $x_0 + \lambda y \in P \cap C$. Hence, $f(x)$ has a finite maximal value f^* on C.

Now assume that f^* can not be attained. Then there exists an unbounded sequence $\{x_n\}_{n=1}^{\infty} \subseteq C$ such that $f(x_n) \geq f^* - 1/n$. Consider the closed convex set

$$C' = \mathbf{cl}\left(\mathbf{co}\left(\{x_n\}_{n=1}^{\infty} \cup \{x_0\}\right)\right),$$

we have $f(x) \geq \min\{f^* - 1, f(x_0)\}$ on C'. Since C' is unbounded and $x_0 \in \mathbf{ri}(C)$, there exists a non-zero vector $y \in 0^+(C') \subseteq 0^+C$ by Theorem 8.4 in [2] and Proposition 2.1. Then $x_0 + \lambda y \in C'$ for all $\lambda \geq 0$ and $f(y) < 0$ since $a \in \mathbf{int}\left((0^+C)^o\right)$. Therefore, f is unbounded from below on C', which is a contradiction.

If $a \notin (0^+C)^o$, then there exists a vector $y \in 0^+C$ such that $x_0 + \lambda y \in C$ for all $\lambda \geq 0$ and $f(y) = a^T y > 0$. Therefore $f(x_0 + \lambda y)$ is unbounded and the maximal value of $f(x) = a^T x$ on C is infinite. Hence, we have $\mathbf{dom}\left(\delta^*(x^* \mid C)\right) \subseteq (0^+C)^o$. $\square$

Remark 2.3. *For a vector $a \in (0^+C)^o \backslash \mathbf{int}\left((0^+C)^o\right)$, it is difficult to determine whether $a \in \mathbf{dom}\left(\delta^*(x^* \mid C)\right)$. Moreover, the maximal value of $f(x) = a^T x$ could still be unattainable even when it is finite.*

Example 2.4. *Let us consider a closed convex set C defined by*

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 \geq 1,\ x \geq 0\}.$$

We have

$$0^+C = \{(x, y) \in \mathbb{R}^2 \mid -x \leq y \leq x,\ x \geq 0\},$$

and

$$(0^+C)^o = \{(x, y) \in \mathbb{R}^2 \mid x \leq y \leq -x,\ x \leq 0\}.$$

Hence $(-1, 1)$ is on the boundary of $(0^+C)^o$ and it corresponds to a linear function $f(x) = -x + y$ with maximal value 0 which is not attainable.

Let us review some background about dual variety in projective space $\mathbb{P}^n$ [3, 4]. Let $I = \langle f_1, \ldots, f_p \rangle$ be a homogeneous radical ideal in polynomial ring $\mathbb{R}[x_0, x_1, \ldots, x_n]$ and $X = V(I)$ be the variety of I in $\mathbb{P}^n$ over $\mathbb{C}$. The *singular locus* $\mathrm{sing}(X)$ is defined by the vanishing of the $c \times c$ minors of the $p \times (n+1)$ Jacobian matrix $\mathrm{Jac}(X) = (\partial f_i / \partial x_j)$, where $c = \mathrm{codim}(X)$. Let $X_{\mathrm{reg}} = X \backslash \mathrm{sing}(X)$ denote the set of regular points in X. The projective variety X is smooth if $X = X_{\mathrm{reg}}$. A point $(u_0 : u_1 : \cdots : u_n)$ in the dual projective space $(\mathbb{P}^n)^*$ represents the hyperplane $\{x \in \mathbb{P}^n \mid \sum_{i=0}^n u_i x_i = 0\}$. The vector u is *tangent* to X at a regular point $x \in X_{\mathrm{reg}}$ if x lies in the hyperplane and its representing vector $(u_0, u_1, \ldots, u_n)$ lies in the row space of the Jacobian matrix $\mathrm{Jac}(X)$ at the point x. The *conormal variety* $\mathrm{CN}(X)$ is the closure of the set

$$\{(x, u) \in \mathbb{P}^n \times (\mathbb{P}^n)^* \mid x \in X_{\mathrm{reg}} \text{ and } u \text{ is tangent to } X \text{ at } x\}.$$

The *dual variety* X^* is the projection of $\mathrm{CN}(X)$ onto the second factor.

Theorem 2.5. *Let $X^* \subset (\mathbb{P}^n)^*$ be the dual variety to the projective closure of a real affine variety X in $\mathbb{R}^n$ and $C = \mathbf{cl}\left(\mathbf{co}(X)\right)$. If X is irreducible, smooth and 0^+C is pointed, then X^* is an irreducible hypersurface, and its defining polynomial is $\Phi(-c_0, c_1, \ldots, c_n)$, where Φ represents the optimal value function.*

Proof. Fix a vector $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$ such that the maximal value c_0^* of $f(x) = c^T x$ on X is finite. We first show that $(-c_0^* : c_1 : \cdots : c_n)$ belongs to X^*.

Case I: c_0^* is attainable. Let the maximal value c_0^* be reached at $x^* \in X$. Since X is smooth, x^* is a regular point of X and KKT conditions are satisfied at x^*. Then the argument used in the proof of Theorem 5.23 in [4] is still valid, which implies that $(-c_0^* : c_1 : \cdots : c_n)$ lies in the dual variety X^*.

Case II: c_0^* is an asymptotic maximal value, i.e., c_0^* is only reached at infinity. Since 0^+C is pointed, by Theorem 2.2 and Remark 2.3, $(0^+C)^o$ is n-dimensional and $c \in (0^+C)^o \backslash \mathbf{int}\left((0^+C)^o\right)$. Fix a vector $a \in \mathbf{int}\left((0^+C)^o\right)$. Since $\delta^*(x^* \mid C)$ is a proper convex function, by Theorem 7.5 in [2], we have

$$c_0^* = \delta^*(c \mid C) = \lim_{\lambda \to 0} \delta^*((1 - \lambda)c + \lambda a \mid C).$$

Then $(c_0^*, c) = \lim_{\lambda \to 0} (\delta^*((1-\lambda)c + \lambda a \mid C), (1-\lambda)c + \lambda a)$. By Theorem 6.1 in [2], we have $(1 - \lambda)c + \lambda a \in \mathbf{int}\left((0^+C)^o\right)$ for $0 < \lambda \leq 1$. By Theorem 2.2, the maximal value of the function $((1 - \lambda)c + \lambda a)^T x$ on X is finite and attainable for $0 < \lambda \leq 1$. Using the results in Case I, we have $(-\delta^*((1 - \lambda)c + \lambda a \mid C) : (1 - \lambda)c + \lambda a)$ lies in X^* for $0 < \lambda \leq 1$, which implies $(-c_0^* : c_1 : \cdots : c_n) \in X^*$ since X^* is closed.

We claim that X^* is a hypersurface. In fact, for every $c \in \mathbf{int}\left((0^+C)^o\right)$, by Theorem 2.2, the function $c^T x$ has a finite maximal value c_0^* on X and therefore $(-c_0^* : c_1 : \cdots : c_n) \in X^*$ by the results in Case I. Since $(0^+C)^o$ is n-dimensional, we deduce that X^* is of dimension $n - 1$. As X is irreducible, X^* is also irreducible and hence it is a hypersurface. The defining equation of X^* can be written as $\Phi(-c_0, c_1, \ldots, c_n) = 0$, then Φ represents the optimal value function of the least degree. Any polynomials containing the graph of the optimal value function must be divided by Φ. $\square$

3. DISCUSSION

Consider the case when X is noncompact and 0^+C is unpointed. By definition, 0^+C contains a line $L \subset \mathbb{R}^n$ through the origin and therefore $(0^+C)^o$ is a subset of the hyperplane $H \subset \mathbb{R}^n$ which is orthogonal to L. By Theorem 2.2 (b), we have $\mathbf{dom}\left(\delta^*(x^* \mid C)\right) \subseteq H$. Since the dual variety X^* is typically expected to have dimension $n - 1$ (see [3, 4]), it is perhaps not the closure of

$$\{(-c_0^* : c_1 : \cdots : c_n) \in (\mathbb{P}^n)^* \mid c \in \mathbf{dom}\left(\delta^*(x^* \mid C)\right)\}. \quad (3.1)$$

However, we still can prove (3.1) belongs to X^*.

4. REFERENCES

[1] J. M. Borwein and A. S. Lewis. *Convex Analysis and Nonlinear Optimization*. CMS Books in Mathematics. Springer, 2006.

[2] R. T. Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, Princeton, 1996.

[3] P. Rostalski and B. Sturmfels. Dualities in convex algebraic geometry. *Rendiconti di Matematica, Serie VII*, 30:285–327, 2010.

[4] P. Rostalski and B. Sturmfels. Dualities. In G. Blekherman, P. A. Parrilo, and R. R. Thomas, editors, *Semidefinite Optimization and Convex Algebraic Geometry*, MOS-SIAM Series on Optimization, chapter 5, 203–250. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2012.