A PRACTICAL ALGORITHM FOR MINIMIZING A RANK-TWO SADDLE FUNCTION ON A POLYTOPE

Takahito Kuno
University of Tsukuba

(Received February 25, 1994; Final December 5, 1994)

Abstract  This paper addresses a practical method for minimizing a class of saddle functions $f : \mathbb{R}^n \to \mathbb{R}$ on a polytope. Function $f$ is continuous and possesses a rank-two property, i.e., the value of $f$ is defined only by two linearly independent vectors. It is shown that a parametric right-hand-side simplex algorithm decomposes the problem into a finite sequence of one-dimensional subproblems. A globally $\epsilon$-optimal solution of each subproblem is obtained by using a successive underestimation method. Computational results indicate that the algorithm can solve fairly large scale problems efficiently.

1. Introduction

In this paper we will develop a practical algorithm for minimizing a class of saddle functions $f : \mathbb{R}^n \to \mathbb{R}$, i.e.,

$$\min \{ f(x) \mid x \in D \},$$

where $D \subset \mathbb{R}^n$ is a polytope. We assume that $f$ is continuous and possesses the rank-two property with respect to two linearly independent vectors $c_1, c_2 \in \mathbb{R}^n$. This means that there exists a continuous function $g : \mathbb{R}^2 \to \mathbb{R}$ such that $f(x) = g(c_1^T x, c_2^T x)$ for all $x \in \mathbb{R}^n$ [15], though we need not know $g$ explicitly in our algorithm. Since $f$ is a saddle function, $g(\cdot, c_2^T x)$ and $g(c_1^T x, \cdot)$ are convex and (quasi)concave functions respectively for any fixed $x \in \mathbb{R}^n$. Due to this convex-concave property of $f$, there are multiple locally optimal solutions in $D$. In contrast to (quasi)concave minimization problems, (1.1) might have no globally optimal solutions among vertices of $D$.

Saddle functions are well known in many literature in the context of minimax problems. In [17] Muu and Oettli have solved a more general class of (1.1), in which $f$ is a full-rank saddle function. Muu has also considered a problem containing a full-rank saddle function in the constraint rather than in the objective function [16]. However, the algorithms developed for the general purpose can usually handle only instances of a very limited scale. We will therefore exploit the rank-two property of $f$ and show that a parametric simplex algorithm decomposes (1.1) into a finite sequence of one-dimensional subproblems, which can be solved very efficiently.

Rank-two nonconvex minimization problems are important in practical applications such as bicriterion decision making [4, 8], computational geometry [11, 14] or network flow problems [25] to name only a few (see [24]). Many of them, involving linear multiplicative programs [9, 18, 23] and certain d.c. programs (minimizations of the difference of two convex functions $h_1(c_1^T x) - h_2(c_2^T x)$) [22], belong to the class (1.1).

The author was partially supported by Grand-in-Aid for Scientific Research of the Ministry of Education, Science and Culture, Grant No. (C)05650061.
In Section 2 we will show that (1.1) can be solved by solving a sequence of one-dimensional problems of the same form as (1.1). The sequence can be generated by applying a parametric right-hand-side simplex algorithm to two linear programs associated with (1.1). Section 3 is devoted to the procedure for obtaining a globally $\varepsilon$-optimal solution of one-dimensional problems. By exploiting the convex-concave property of $f$ we will construct a branch-and-bound algorithm based on a successive underestimation method [7]. Results of computational experiment on the algorithm are presented in Section 4. In Section 5 we will briefly discuss the average performance of the algorithm when we apply it to certain nonconvex quadratic programs.

2. Decomposition of the Problem into One-Dimensional Problems

The problem we consider in this paper is as follows:

\[
\begin{align*}
\text{(P)} & \quad \text{minimize} \quad f(x) \\
& \quad \text{subject to} \quad Ax = b, \ x \geq 0,
\end{align*}
\]

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $f : \mathbb{R}^n \to \mathbb{R}^1$ is a continuous function. There are two linearly independent vectors $c_1, c_2 \in \mathbb{R}^n$ which characterize $f$. Namely,

(i) Rank-two property: For any $x \in \mathbb{R}^n$

\[d \in \mathbb{R}^n, \ c_k^T d = 0, \ k = 1, 2 \Rightarrow f(x + d) = f(x).\]  

(ii) Convex-concave property: For any $x \in \mathbb{R}^n$

\[
\begin{align*}
d \in \mathbb{R}^n, \ c_2^T d = 0 & \Rightarrow f(x + \lambda d) \leq (1 - \lambda)f(x) + \lambda f(x + d), \ \forall \lambda \in [0, 1], \\
d \in \mathbb{R}^n, \ c_1^T d = 0 & \Rightarrow f(x + \lambda d) \geq \min\{f(x), f(x + d)\}, \ \forall \lambda \in [0, 1].
\end{align*}
\]

We assume in the sequel that the feasible region:

\[D = \{x \in \mathbb{R}^n \mid Ax = b, \ x \geq 0\}.
\]

is nonempty and bounded, which implies that (P) has a globally optimal solution. Figure 2.1 shows a two-dimensional example, where $c_1 = (0, 1)^T$, $c_2 = (1, 0)^T$ and $f(x) = -x_1^2 + x_2^2$. It is easy to see that this function has three local minimum points A, B and C, among which C is the global one.

Let $\zeta = c_1^T x$ for an arbitrary $x \in D$ and consider a subproblem of (P):

\[
\begin{align*}
\text{(P(\zeta))} & \quad \text{minimize} \quad f(x) \\
& \quad \text{subject to} \quad x \in D, \ c_1^T x = \zeta.
\end{align*}
\]

Then (P(\zeta)) is feasible and has an optimal solution which coincides with that of a linear program, i.e., either

\[
\begin{align*}
\text{(PL}_1(\zeta)) & \quad \text{minimize} \quad c_2^T x \\
& \quad \text{subject to} \quad x \in D, \ c_1^T x = \zeta,
\end{align*}
\]

or

\[
\begin{align*}
\text{(PL}_2(\zeta)) & \quad \text{maximize} \quad c_2^T x \\
& \quad \text{subject to} \quad x \in D, \ c_1^T x = \zeta.
\end{align*}
\]

Let $x^k(\zeta)$ be an optimal solution of (PL$k(\zeta)) (k = 1, 2$) and define

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.
Lemma 2.1. If $\zeta = c_1^T x$ for some $x \in D$, then $x^*(\zeta)$ is optimal to (P($\zeta$)).

Proof: By the rank-two property, $f$ is a function of a single variable $\eta = c_2^T x$ if the value $c_1^T x$ is fixed at $\zeta$. The values $c_2^T x^1(\zeta)$ and $c_2^T x^2(\zeta)$ are the minimum and the maximum of $\eta$ respectively. It follows from (2.3) of property (ii) that the minimum of $f$ is attained at either of the extreme points of the interval $[c_2^T x^1(\zeta), c_2^T x^2(\zeta)]$. 

Let

$$
\zeta_{\min} = \min\{c_1^T x \mid x \in D\}; \quad \zeta_{\max} = \max\{c_1^T x \mid x \in D\}.
$$

It is obvious that a globally optimal solution of (P) can be obtained by solving (P($\zeta$)) for all $\zeta \in [\zeta_{\min}, \zeta_{\max}]$. By Lemma 2.1, this can be done if we solve the two linear programs (PL$_1(\zeta)$) and (PL$_2(\zeta)$) as varying the value $\zeta$ over the interval $[\zeta_{\min}, \zeta_{\max}]$.

Theorem 2.2. There exists $\zeta \in [\zeta_{\min}, \zeta_{\max}]$ such that $x^*(\zeta)$ is a globally optimal solution of (P).

Let us apply a parametric right-hand-side simplex method (abbreviated as PRSM) to (PL$_k(\zeta)$) ($k = 1, 2$). For the sake of simplicity we impose here the dual nondegeneracy assumption:

Assumption 2.1. Both (PL$_1(\zeta)$) and (PL$_2(\zeta)$) have a unique optimal solution for any $\zeta \in [\zeta_{\min}, \zeta_{\max}]$. 

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.
Suppose we have an optimal basis $B^*_0 \in \mathbb{R}^{(m+1) \times (m+1)}$ of $(PL_k(\zeta_{\text{min}}))$. Under the above assumption $B^*_0$ remains optimal even if the value of $\zeta$ slightly increases from $\zeta_{\text{min}}$. However, when $\zeta$ is beyond some point, say $\zeta^+_k (\leq \zeta_{\text{max}})$, some basic variable turns negative and the primal feasibility of $B^*_0$ is violated. Then we carry out a single dual pivot replacing the basic variable with an appropriate nonbasic one and obtain an alternative basis $B^*_1$, which is optimal to $(PL_k(\zeta^+_k))$ (see e.g. [2, 3] for further details).

Applying these operations iteratively, we can generate a sequence of subintervals $[\zeta^+_0, \zeta^+_1], [\zeta^+_1, \zeta^+_2], \ldots, [\zeta^+_k, \zeta^+_{k+1}]$ in the interval $[\zeta_{\text{min}}, \zeta_{\text{max}}]$, where $\zeta^+_0 = \zeta_{\text{min}}, \zeta^+_k = \zeta_{\text{max}}$ and $\zeta^+_{k+1} > \zeta^+_k$ for each $i$. Simultaneously, we have the associated sequence of bases $B^*_0, B^*_1, \ldots, B^*_{p_k-1} \in \mathbb{R}^{(m+1) \times (m+1)}$ such that $B^*_k$ is optimal to $(PL_k(\zeta))$ for all $\zeta \in [\zeta^+_k, \zeta^+_{k+1}]$. We denote $[\zeta^+_k, \zeta^+_{k+1}]$ by $Z^+_k$ in the sequel. As well known, $x^k(\zeta)$ is an affine function over each $Z^+_k$ and can be expressed as

$$x^k(\zeta) = \frac{\zeta_{i+1} - \zeta}{\zeta_{i+1} - \zeta_i} x^k(\zeta_i^k) + \frac{\zeta - \zeta_i}{\zeta_{i+1} - \zeta_i} x^k(\zeta_{i+1}^k), \ \zeta \in Z^+_i.$$ 

If for every $i$ we can compute

$$x^k(Z^+_i) \in \text{argmin}\{f(x) \mid x = (1 - \lambda)x^k(\zeta_i^k) + \lambda x^k(\zeta_{i+1}^k), \ \lambda \in [0, 1]\},$$

then Theorem 2.2 guarantees that

$$x^* \in \text{argmin}\{f(x) \mid x = x^k(Z^+_i), \ j = 0, 1, \ldots, p_k - 1, \ k = 1, 2\}$$

is a globally optimal solution of (P). The procedure for computing $x^k(Z^+_i)$ will be presented in the next section.

We summarize the algorithm below:

**Algorithm PRSM.**

**Step 1.** Solve a linear program: minimize $\{c_1^T x \mid x \in D\}$ and obtain an optimal basis $B^*$ and the associated optimal solution $x^*$. Initialize the incumbent: $x^* = x^0, v^* = f(x^*)$. Let $k = 1$ and go to Step 2.

**Step 2.** Let $\zeta = c_1^T x^*$ and $B = B^*$. Solve a linear program $(PL_k(\zeta))$ parametrically by increasing $\zeta$ from $\zeta$:

1° If $(PL_k(\zeta))$ is infeasible for $\zeta > \zeta$, then go to Step 3.

2° Determine a value $\bar{\zeta}$ of $\zeta$ such that $B$ is an optimal basis for all $\zeta \in Z = [\zeta, \bar{\zeta}]$.

Using a dual pivot operation, obtain an alternative basis $\bar{B}$ which is optimal to $(PL_k(\bar{\zeta}))$.

3° Compute $x^k(Z) \in \text{argmin}\{f(x) \mid x = (1 - \lambda)x^k(\bar{\zeta}) + \lambda x^k(\zeta), \ \lambda \in [0, 1]\}$. If $f(x^k(Z)) < v^*$, then update the incumbent: $x^* = x^k(Z), v^* = f(x^*)$.

4° Let $\zeta = \bar{\zeta}, B = \bar{B}$ and go to 1°.

**Step 3.** If $k = 2$, then terminate. Otherwise, let $k = 2$ and go to Step 2.

Under Assumption 2.1, the above algorithm terminates after finitely many iterations yielding an optimal solution $x^*$ of (P) if Step 2. 3° can be done in finite time. In the case of degeneracy, we have to use a suitable pivoting rule to avoid cycling (see e.g. [2]).
3. Successive Underestimation Method for One-Dimensional Problems

In this section we consider the problem to be solved in Step 2, 3° of algorithm PRSM, i.e., for each $k = 1, 2,$

$$(P_k(Z)) \begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad x = (1 - \lambda)x^k(\zeta) + \lambda x^k(\bar{\zeta}). \quad \lambda \in [0, 1],
\end{align*}$$

where $x^k(\zeta)$ and $x^k(\bar{\zeta})$ are optimal solutions of $(PL_k(\zeta))$ and $(PL_k(\bar{\zeta}))$ respectively, and $Z = [\zeta, \bar{\zeta}]$ is a subinterval of $[\zeta_{\text{min}}, \zeta_{\text{max}}]$ such that a basis $\mathcal{B}$ is optimal to $(PL_k(\zeta))$ for all $\zeta \in Z$. The difference between $(P_k(Z))$ and $(P)$ is that the feasible region of the former:

$$D_k(Z) = \{x \in \mathbb{R}^n \mid x = (1 - \lambda)x^k(\zeta) + \lambda x^k(\bar{\zeta}), \quad \lambda \in [0, 1]\}$$

is only a line segment. Hence, if $f$ is either convex or concave over $D_k(Z)$, we can compute a minimum $x^k(Z)$ very efficiently by using any one of ordinary methods. This involves the case in which either $c_2^T x$ or $c_1^T x$ is a constant for any $x \in D_k(Z)$. Although both the values are affine functions of $\lambda$ over $D_k(Z)$, they are not constants in general. We will therefore propose a successive underestimation method for obtaining a globally $\epsilon$-optimal solution of $(P_k(Z))$.

3.1. LOWER BOUNDS OF THE OBJECTIVE FUNCTION VALUE

We first define a vector $\tilde{c}_1 \in \mathbb{R}^n$ below:

$$\tilde{c}_1 = c_1 - (c_1^T c_2 / \|c_2\|^2) c_2. \quad (3.1)$$

Then we have $c_1^T \tilde{c}_1 > 0$ and $c_2^T \tilde{c}_1 = 0$ by noting that $c_1$ and $c_2$ are linearly independent. Hence by (2.2) of property (ii) function $f$ is convex with respect to the direction $\tilde{c}_1$. We can compute the following by using convex minimization:

$$L_k(Z) = \arg\min \{f(x) \mid x = x^k(\zeta) + \alpha(Z) \tilde{c}_1, \quad \lambda \in [0, 1]\}, \quad (3.2)$$

$$L_k(Z) = \arg\min \{f(x) \mid x = x^k(\bar{\zeta}) - \alpha(Z) \tilde{c}_1, \quad \lambda \in [0, 1]\}. \quad (3.3)$$

where

$$\alpha(Z) = (\bar{\zeta} - \zeta) / (c_1^T \tilde{c}_1). \quad (3.4)$$

Let

$$v^k(Z) = \min \{f(x) \mid x \in L_k(Z) \cup L_k(Z)\}. \quad (3.5)$$

**Lemma 3.1.** For any subinterval $Z' = [\zeta', \bar{\zeta}'] \subset Z$ the following relationship holds:

$$v^k(Z) \leq v^k(Z').$$

*Proof:* Choose an arbitrary $x' \in L_k(Z')$. Then there exists $\lambda' \in [0, 1]$ such that $x' = x^k(\zeta') + \lambda' \alpha(Z') \tilde{c}_1$. By linearity of $x^k(\zeta)$ over $Z$ and definition of $\alpha(Z)$ we have

$$x^k(\zeta') = \frac{\zeta' - \zeta}{\bar{\zeta} - \zeta} x^k(\bar{\zeta}) + \frac{\zeta' - \zeta}{\zeta - \zeta} x^k(\zeta), \quad \alpha(Z') = \frac{\bar{\zeta} - \zeta'}{\zeta - \zeta} \alpha(Z).$$

Hence $x'$ is written as

$$x' = (1 - \beta)x^k(\zeta) + \beta x^k(\bar{\zeta}) + \lambda' \alpha(Z) \tilde{c}_1. \quad (3.6)$$
where \( \beta = (\zeta' - \zeta) / (\zeta - \zeta) \) and \( \gamma = (\zeta' - \zeta) / (\zeta - \zeta) \). Let
\[
\underline{x} = x^k(\zeta) + (\beta + \lambda')\alpha(Z)\bar{c}_1; \quad \bar{x} = x^k(\bar{\zeta}) - (1 - \beta - \lambda')\alpha(Z)\bar{c}_1.
\]
Then (3.6) is reduced to the following:
\[
x' = (1 - \beta)\underline{x} + \beta\bar{x}.
\]
Since \( c^T_1x^k(\zeta) = \zeta \) and \( c^T_1x^k(\bar{\zeta}) = \bar{\zeta} \) by definition, we see that
\[
c^T_1(\bar{x} - \underline{x}) = c^T_1x^k(\bar{\zeta}) - c^T_1x^k(\zeta) - \alpha(Z)c^T_1\bar{c}_1 = \bar{\zeta} - \zeta - (\zeta - \zeta) = 0
\]
by noting (3.4). We can also check that \( \beta + \gamma \lambda' \in [0, 1] \) if \( Z' \subset Z \). Hence by (2.3) of property (ii) and definition of \( v^k(Z) \) we obtain
\[
f(x') \geq \min \{f(\underline{x}), f(\bar{x})\} \geq v^k(Z).
\]
Similarly, we have \( f(x) \geq v^k(Z) \) for any \( x \in \bar{L}_k(Z') \). \( \square \)

As a corollary of this lemma, we can show that \( v^k(Z) \) gives a lower bound of the optimal value \( f(x^k(Z)) \) of problem \( (P_k(Z)) \):

**Lemma 3.2.** For any \( x \in D_k(Z) \) the following holds:
\[
f(x) \geq v^k(Z). \tag{3.7}
\]

**Proof:** For any \( x \in D_k(Z) \) there exists some \( \zeta' \in Z \) such that \( x = x^k(\zeta') \). Hence (3.7) is derived by applying Lemma 3.1 to \( Z' = [\zeta', \zeta] \subset Z \). \( \square \)

Note that \( v^k(Z) \) is the optimal value of a relaxed problem of \( (P_k(Z)) \):

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad x \in R (\supset D_k(Z)),
\end{align*}
\]

where \( R \) is a rectangle with vertices \( x^k(\zeta) \) and \( x^k(\bar{\zeta}) \) in the plane spanned by \( c_1 \) and \( c_2 \). Each side of \( R \) is collinear with either \( \bar{c}_1 \) or \( c_2 \), and hence by quasiconcavity of \( f \) with respect to \( c_2 \) the minimum is achieved at some points, say \( L_k(Z) \) or \( \bar{L}_k(Z) \), on the sides collinear with \( \bar{c}_1 \).

Lemma 3.2 enables us to discard \( D_k(Z) \) in the course of locating a globally optimal solution of \( (P) \) in PRSM when
\[
v^k(Z) \geq f(x^*) \tag{3.8}
\]
holds for the best feasible solution \( x^* \) obtained by that time. In this case we cannot update the incumbent better than \( x^* \) by any point of \( D_k(Z) \).

### 3.2. Branch-and-Bound Procedure

Let us suppose that (3.8) does not hold. When some point \( x' \) in the set:
\[
L_k(Z) = \{ x \in L_k^k(Z) \cup \bar{L}_k^k(Z) \mid f(x) = v^k(Z) \}
\]
is found to be a feasible solution of \( (P) \), we may discard \( D_k(Z) \) and proceed to the next step after revising the incumbent \( x^* \) by \( x' \). If such an \( x' \) cannot be found, i.e., \( L_k(Z) \cap D = \emptyset \),
we have to search $D_k(Z)$ for a better feasible solution than $x^*$.

Let us bisect the interval $Z = [\zeta, \bar{\zeta}]$ into $Z_{11} = [\zeta, \zeta_0]$ and $Z_{12} = [\zeta_0, \bar{\zeta}]$, where $\zeta_0 = (\zeta + \bar{\zeta}) / 2$. Then the value $f(x^k(\zeta_0))$ is an upper bound of the optimal value of $(P_k(Z))$. If $f(x^k(\zeta_0)) < f(x^*)$, then we need to update the incumbent as $x^* = x^k(\zeta_0)$. Note that we can compute $x^k(\zeta_0)$ without performing any pivot operations, since $x^k(\zeta)$ is affine over the interval $Z$. We next construct the problems $(P_k(Z_{11}))$ and $(P_k(Z_{12}))$ associated with the intervals $Z_{11}$ and $Z_{12}$ respectively, and compute lower bounds $v^k(Z_{11})$ and $v^k(Z_{12})$ of their optimal values. It is obvious that $D_k(Z_{11}) \cup D_k(Z_{12}) = D_k(Z)$ and $D_k(Z_{11}) \cap D_k(Z_{12}) = \{x^k(\zeta_0)\}$. Let us define a piecewise constant function on $D_k(Z)$:

$$g_1(x) = \begin{cases} v^k(Z_{11}), & x \in D_k(Z_{11}), \\ v^k(Z_{12}), & x \in D_k(Z) \setminus D_k(Z_{11}). \end{cases}$$

Then by Lemma 3.1 we see that

$$v^k(Z) \leq g_1(x) \leq f(x), \ \forall x \in D_k(Z).$$

A further bisection of $Z_{12}$ with $v^k(Z_{12}) = \min\{v^k(Z_{11}), v^k(Z_{12})\}$ at its middle point $\zeta_1$ can generate an alternative function $g_2$, which underestimates $f$ over $D_k(Z)$ more exactly than $g_1$.

If we iterate the above operations as selecting one subinterval of $Z$ giving the least lower bound among them, we will obtain a sequence of piecewise constant functions $g_j$'s such that

$$(v^k(Z) = g_0(x) \leq g_1(x) \leq g_2(x) \leq \cdots \leq f(x), \ \forall x \in D_k(Z).$$

Note that $x^k(\zeta_j)$ is a minimizer of $g_j$ and a jumping point of $g_{j+1}$. The incumbent $x^*$ is updated by $x^k(\zeta_j)$ when necessary.

If $g_j(x^k(\zeta_j)) \geq f(x^*)$ happens to hold, then two cases are possible: (i) $x^*$ is an optimal solution of $(P_k(Z))$ if $x^* \in D_k(Z)$, (ii) there are no globally optimal solution of $(P)$ in $D_k(Z)$ otherwise. In either case we can terminate the procedure. Figure 3.1 illustrates the procedure when we apply it to the example shown in Section 2. Here $D_k(Z)$ corresponds to the edge B-C of $D$ in Figure 2.1.

The procedure is summarized as the following branch-and-bound algorithm. Here $\epsilon \geq 0$ is a given tolerance, $x^*$ and $v^*$ are the incumbent and its objective function value respectively.

**Procedure BBP**$(k, x^*, v^*, Z)$.

1. Compute $v^k(Z)$ and $L_k(Z)$ according to (3.1) – (3.5) and (3.9). If $v^k(Z) \geq f(x^*)$, then terminate. Otherwise, let $\mathcal{Z} = \{Z\}$ and $j = 0$.

2. Select an interval $Z_j = [\zeta_j, \bar{\zeta}_j] \in \mathcal{Z}$ with the least $v^k(Z_j)$ and let $\mathcal{Z} = \mathcal{Z} \setminus \{Z_j\}$. If $L_k(Z_j) \cap D \neq \emptyset$, then terminate after revising the incumbent: $x^* = x'$, $v^* = f(x^*)$ for an arbitrary $x' \in L_k(Z_j) \cap D$.

3. Let $\xi_j = (\zeta_j + \bar{\zeta}_j) / 2$. If $f(x^k(\zeta_j)) < v^*$, then update the incumbent: $x^* = x^k(\zeta_j)$, $v^* = f(x^*)$. If

$$f(x^*) - v^k(Z_j) \leq \epsilon,$$  \hspace{1cm} (3.10)
Figure 3.1. Illustration of BBF.

then terminate.

4° Let $Z_j = [\zeta_j, \zeta_j]$ and $\overline{Z}_j = [\zeta_j, \overline{\zeta}_j]$. Compute $v^k(Z_j)$, $L_k(Z_j)$, $v^k(\overline{Z}_j)$ and $L_k(\overline{Z}_j)$.

5° Let $\mathcal{Z} = \mathcal{Z} \cup \{Z_j, \overline{Z}_j\}$. Let $j = j + 1$, and go to 2°.

Theorem 3.3. Procedure BBP terminates after finitely many iterations if $\epsilon > 0$. If $\epsilon = 0$ and BBP does not terminate, it generates an infinite sequence of points $x^k(\zeta_j)$'s, every accumulation point of which is a globally optimal solution of $(P_k(Z))$.

Proof: Suppose the procedure does not terminate. Then an infinite sequence of intervals $Z_j = [\zeta_j, \zeta_j]$'s is generated in $Z$. We can take a subsequence $Z_{j_\ell}$'s such that $(Z =) Z_{j_0} \supset Z_{j_1} \supset Z_{j_2} \supset \cdots$. Since $Z_j$ is divided by the middle point $\zeta_j = (\zeta_j + \overline{\zeta}_j) / 2$, we can assume that $\overline{\zeta}_{j_\ell} - \zeta_{j_\ell} = 2(\overline{\zeta}_{j_{\ell+1}} - \zeta_{j_{\ell+1}})$ for every $\ell$. Hence we have

$$\|x^k(\overline{\zeta}_{j_\ell}) - x^k(\zeta_{j_\ell})\| = \|x^k(\zeta_j) - x^k(\overline{\zeta}_j)\| / 2^\ell$$

(3.11)

by linearity of $x^k(\zeta)$ over $D_k(Z)$.

Now we assume that there exists some positive constant $\sigma$ such that

$$f(x^*) - v^k(Z_{j_\ell}) \geq \sigma, \quad \forall \ell.$$  

(3.12)

By continuity of $f$ there is some positive value $\delta(\sigma)$ such that if

$$\|x' - x''\| < \delta(\sigma),$$

(3.13)

then $|f(x') - f(x'')| < \sigma$. It follows from (3.11) that (3.13) holds for any $x', x'' \in D_k(Z_{j_\ell})$ when $\ell$ is beyond a number:
\[ \ell(\sigma) = \ln \|x^k(\zeta) - x^k(\zeta)\| - \ln \delta(\sigma). \]

Moreover, we can see from (3.2) - (3.5) and (3.9) that if \( x' \in L_k(Z_{ji}) \), i.e., \( f(x') = \ell^k(Z_{ji}) \), then \( \|x' - x\| < \delta(\sigma) \) for any \( x \in D_k(Z_{ji}) \). Therefore we have \( f(x^k(\zeta_{ji})) - \ell^k(Z_{ji}) < \sigma \) for \( \ell > \ell(\sigma) \), which contradicts assumption (3.12). If \( \epsilon > 0 \), then (3.10) holds after finitely many iterations and BBP terminates.

Suppose \( \epsilon = 0 \). Then we have \( \lim_{\ell \to \infty} (f(x^k(\zeta_{ji})) - \ell^k(Z_{ji})) = 0 \). Since we choose \( Z_{ji} \) with the least \( \ell^k(Z_{ji}) \) from \( Z_j \) we obtain

\[ \lim_{\ell \to \infty} f(x^k(\zeta_{ji})) = \lim_{\ell \to \infty} \ell^k(Z_{ji}) \leq f(x), \forall x \in D_k(Z). \]

To save the memory needed by BBP we can employ the depth first rule in choosing \( Z_j \) from \( Z \) instead of the best bound rule. Although the convergence is somewhat slower, this modification causes no trouble if \( \epsilon > 0 \). However, if \( \epsilon = 0 \), the sequence \( x^k(\zeta_j) \)'s might converge to some locally but not globally optimal solution of \( (P_k(Z)) \).

4. Computational Experiment
We will report the results of computational experiment on algorithm PRSM incorporating procedure BBP. We solved the following two subclasses of \( (P) \):

\[
\begin{align*}
\text{minimize} & \quad (c_1^T x - c_{10})^2 - (c_1^T x - c_{10})(c_2^T x - c_{20}) \\
\text{subject to} & \quad Ax \leq b, \quad x \geq 0, \\
& \quad c_k^T x \geq c_{k0}, \quad k = 1, 2, \\
\end{align*}
\]

\[
\begin{align*}
\text{minimize} & \quad (c_1^T x - c_{10})^2 - (c_1^T x - c_{10}) \exp(c_{20} - c_2^T x) \\
\text{subject to} & \quad Ax \leq b, \quad x \geq 0, \\
& \quad c_k^T x \geq c_{k0}, \quad k = 1, 2, \\
\end{align*}
\]

where \( c_k \in \mathbb{R}^n \), \( c_{k0} \in \mathbb{R} \), \( A \in \mathbb{R}^{m \times n} \), and \( b \in \mathbb{R}^m \). All data of examples were randomly generated between \(-1.000 \) and \( 1.000 \). Problem (4.1) is a so-called linear multiplicative program, whose objective function can be expressed by the product of two affine functions, say \( c_1^T x - c_{10} \) and \( (c_1 - c_2)^T x - c_{10} + c_{20} \). If the product is quasiconcave on the feasible region, we can solve the problem efficiently by using the algorithms proposed in [9, 10, 13]. Unfortunately, the objective function of (4.1) is neither convex nor quasiconcave because \( (c_1 - c_2)^T x - c_{10} + c_{20} \) can have both positive and negative values on the feasible region (see e.g. [9]). Hence the available algorithms do not work for (4.1).

In procedure BBP we employed the depth first rule in choosing \( Z_j \) from \( Z \). Also, among two subintervals \( Z_{dj} \) and \( Z_{d_j} \) of \( Z_j \) we took out the one giving the less lower bound from \( Z \) before the other. The program was coded in C language and tested on a SUN SPARCstation ELC computer (20.5 mips).

Table 4.1 shows the computational results when the tolerance is fixed at \( \epsilon = 10^{-5} \) and the size of problems ranges from \((m, n) = (200, 150)\) to \((350, 300)\). It contains the average number of pivot operations (including primal ones for the linear program solved in Step 1 of PRSM), branching operations and the average CPU time in seconds (and also their respective standard deviations in the brackets) needed for solving ten examples. Note that both problems (4.1) and (4.2) require the same number of pivot operations because their feasible regions are identical. Table 4.2 shows the results when \((m, n)\) is fixed at \((200, 150)\) and \( \epsilon \) ranges from \( 10^{-3} \) to \( 10^{-9} \). The average number of branching operations and CPU time of ten examples are listed in it.
Table 4.1. Computational results when $\epsilon = 10^{-5}$.

<table>
<thead>
<tr>
<th>$m$</th>
<th>200</th>
<th>200</th>
<th>250</th>
<th>250</th>
<th>300</th>
<th>300</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>150</td>
<td>200</td>
<td>200</td>
<td>250</td>
<td>250</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>Total number of pivots.</td>
<td>226.4</td>
<td>362.5</td>
<td>385.8</td>
<td>352.5</td>
<td>385.1</td>
<td>463.3</td>
<td>452.1</td>
</tr>
<tr>
<td></td>
<td>(30.124)</td>
<td>(96.225)</td>
<td>(98.238)</td>
<td>(90.155)</td>
<td>(120.45)</td>
<td>(160.515)</td>
<td>(209.555)</td>
</tr>
<tr>
<td>Total number of branchings. (4.1):</td>
<td>138.7</td>
<td>182.8</td>
<td>160.8</td>
<td>152.7</td>
<td>195.9</td>
<td>200.2</td>
<td>153.0</td>
</tr>
<tr>
<td></td>
<td>(78.411)</td>
<td>(100.152)</td>
<td>(83.244)</td>
<td>(98.105)</td>
<td>(82.440)</td>
<td>(122.173)</td>
<td>(102.326)</td>
</tr>
<tr>
<td>(4.2):</td>
<td>151.3</td>
<td>118.2</td>
<td>163.0</td>
<td>189.3</td>
<td>134.6</td>
<td>217.3</td>
<td>145.7</td>
</tr>
<tr>
<td></td>
<td>(82.002)</td>
<td>(94.448)</td>
<td>(94.043)</td>
<td>(140.644)</td>
<td>(94.291)</td>
<td>(106.39)</td>
<td>(121.598)</td>
</tr>
<tr>
<td>CPU time in seconds. (4.1):</td>
<td>46.040</td>
<td>83.130</td>
<td>124.515</td>
<td>117.942</td>
<td>174.678</td>
<td>233.958</td>
<td>279.792</td>
</tr>
<tr>
<td>(4.2):</td>
<td>46.305</td>
<td>82.972</td>
<td>124.525</td>
<td>118.463</td>
<td>173.562</td>
<td>234.020</td>
<td>279.613</td>
</tr>
</tbody>
</table>

We see from Tables 4.1 and 4.2 that algorithm PRSM can solve fairly large scale problems of both the classes (4.1) and (4.2) with enough accuracy when they are randomly generated. There is not much difference in the results between the two classes. It should be noted that the number of branching operations depends only upon the tolerance $\epsilon$ but not upon the size of $(m, n)$. However, since the branching involves no hard operations such as a simplex pivot, it has a little influence on the computational time as shown in Table 4.2. The total computational time is consequently dominated by the number of iterations of the parametric simplex algorithm. Also its variance is reasonably small compared with the usual global optimization algorithms using cutting planes.

5. Average Performance of the Algorithm for Some Instances

As shown in Section 1, problem (P) involves numerous subclasses. Among them are the following two nonconvex quadratic programs:

Table 4.2. Computational results when $(m, n) = (200, 150)$.

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>$10^{-3}$</th>
<th>$10^{-5}$</th>
<th>$10^{-7}$</th>
<th>$10^{-9}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of branchings. (4.1):</td>
<td>36.4</td>
<td>138.7</td>
<td>255.9</td>
<td>373.7</td>
</tr>
<tr>
<td></td>
<td>(24.577)</td>
<td>(78.411)</td>
<td>(139.543)</td>
<td>(203.150)</td>
</tr>
<tr>
<td>(4.2):</td>
<td>34.4</td>
<td>151.3</td>
<td>286.7</td>
<td>423.7</td>
</tr>
<tr>
<td></td>
<td>(18.597)</td>
<td>(82.002)</td>
<td>(141.017)</td>
<td>(200.655)</td>
</tr>
<tr>
<td>CPU time in seconds. (4.1):</td>
<td>45.778</td>
<td>46.040</td>
<td>47.403</td>
<td>47.537</td>
</tr>
<tr>
<td></td>
<td>(6.140)</td>
<td>(5.897)</td>
<td>(6.298)</td>
<td>(6.739)</td>
</tr>
<tr>
<td>(4.2):</td>
<td>45.767</td>
<td>46.305</td>
<td>47.413</td>
<td>47.650</td>
</tr>
</tbody>
</table>
Minimization

Saddle Function

subject to

where \( c_k \in \mathbb{R}^n \), \( c_{k0} \in \mathbb{R}^1 \) (\( k = 1, 2 \)) and \( D \subset \mathbb{R}^n \) defined by (2.4). Linear multiplicative programs (P1) appear in many applications such as microeconomics [6], bond portfolio optimization [8], and computational geometry [11, 14] and so forth (see [10, 18]). If every feasible solution \( x \in D \) satisfies \( c_k^T x \geq c_{k0} \) for \( k = 1, 2 \), then (P1) is a quasiconcave minimization [9] and can be solved by the algorithms proposed in [9, 13] as well. Problem (P2) is a concave quadratic program, whose objective function \( f_2 \) has only one negative eigenvalue. In their recent article [19] Pardalos and Vavasis have proved the NP-hardness of (P2) by converting a clique problem on a graph to (5.2). Quadratic programs are known to be in NP [26], and hence (P2) is a NP-complete problem.

Here we will discuss the average performance of algorithm PRSM when we apply it to those nonconvex quadratic programs (P1) and (P2).

Recall that (\( P_k(Z) \)) solved by BBP is a minimization of \( f \) over the line segment \( D_k(Z) \). If \( f \) is a quadratic function such as \( f_1 \) and \( f_2 \), we can calculate a rigorous solution of (\( P_k(Z) \)) analytically without calling procedure BBP. Hence the total number of arithmetic operations needed for solving (P1) and (P2) can be bounded only by that of dual pivot operations. Moreover, we can solve them even if the feasible region \( D \) is unbounded. In this case the parametric right-hand-side simplex algorithm would generate a basis \( B \) which is optimal to (PL\((k(\zeta)) \) \( k = 1, 2 \)) for all \( \zeta \in \mathbb{Z}' = [\zeta_-, \infty) \) for some \( \zeta_- \). At the same time it generates some direction vector \( d \in \mathbb{R}^n \), and we have

\[
D_k(Z') = \{ x \in \mathbb{R}^n | x = x_k(\zeta) + \lambda d, \lambda \geq 0 \}.
\]

It is easy to check whether \( f_1 \) (\( f_2 \)) is bounded from below on \( D_k(Z') \). If we find it unbounded, the original problem has no globally optimal solutions.

Let us again consider the linear programs (PL\((k(\zeta)) \), \( k = 1, 2 \). Denote by \( g_k(\zeta) \) the objective function value of (PL\((k(\zeta)) \), i.e.,

\[
g_1(\zeta) = \min \{ c_1^T x | x \in D, c_1^T x = \zeta \}; \quad g_2(\zeta) = \max \{ c_2^T x | x \in D, c_1^T x = \zeta \}.
\]

**Lemma 5.1.** Let \( \zeta_- = \inf \{ c_1^T x | x \in D \} \) and \( \zeta_+ = \sup \{ c_1^T x | x \in D \} \). Then, (i) function \( g_1 \) is piecewise linear convex on the interval \( (\zeta_- , \zeta_+) \), (ii) function \( g_2 \) is piecewise linear concave on the interval \( (\zeta_- , \zeta_+) \).

**Proof:** Follows from a well-know result on linear programming (see e.g. [2]).

We can regard PRSM as a method which generates the analytic form of \( g_k \) and compute a global minimum of \( f \) over the line segment corresponding to each linear piece of \( g_k \). Under Assumption 2.1 the number of linear pieces of \( g_k \)'s coincides with that of dual pivot operations of PRSM.

If we take the partial dual with respect to the constraint \( c_1^T x = \zeta \) of (PL\((1(\zeta)) \), then

\[
g_1(\zeta) = \min \{ c_2^T x | x \in D, c_1^T x = \zeta \}
= \min \sup \{ c_2^T x + \eta (c_1^T x - \zeta) | x \in D \}
= \sup \{ -\eta \zeta + \min \{ \eta c_1^T x + c_2^T x | x \in D \} \}.
\]

Letting

\[ Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited. \]
\[
    h_1(\eta) = \min \{ \eta c_1^T x + c_2^T x \mid x \in D \},
\]
we have
\[
    g_1(\zeta) = \sup_{\eta \in \mathbb{R}^1} \{-\eta \zeta + h_1(\eta)\}.
\]
(5.3)

Similarly, \( g_2 \) can be reduced to
\[
    g_2(\zeta) = \inf_{\eta \in \mathbb{R}^1} \{-\eta \zeta + h_2(\eta)\},
\]
where
\[
    h_2(\eta) = \max \{ \eta c_1^T x + c_2^T x \mid x \in D \}.
\]
(5.4)

The following lemma is analogous to Lemma 5.1:

**Lemma 5.2.** (i) Function \( h_1 \) is piecewise linear concave on \( \mathbb{R}^1 \), (ii) function \( h_2 \) is piecewise linear convex on \( \mathbb{R}^1 \). \( \square \)

Thus we can see from (5.3) and (5.4) that if the analytic form of \( h_k \) is given, we can obtain that of \( g_k \) in \( O(I_k) \) time, where \( I_k \) represents the number of linear pieces of \( h_k \). The number of linear pieces of \( g_k \) is obviously \( O(I_k) \).

Adler and Haimovich have proved in [1, 5] that the average number of linear pieces \( I_k \) is bounded by \( O(\min\{m, n\}) \) under sign-invariant probabilistic assumptions imposed on the data \( (A, b, c_1, c_2) \). (Readers are referred to an excellent survey article by Shamir [21] or a book by Schrijver [20] for the results proved in the unpublished manuscripts [1, 5].) In their probabilistic model, Assumption 2.1 is fulfilled with probability one. This implies that the average number of dual pivot operations required by PRSM is also bounded by \( O(\min\{m, n\}) \). On the other hand, the linear program to be solved in Step 1 of PRSM is a standard linear program, which is well known to be solved in polynomial time. Consequently, the average number of arithmetic operations needed for solving (P1) and (P2) is lower-order polynomial relative to the size of \( A \). A similar result for a certain class of bilinear programs has been proved in [12].

The key of the above discussion is the polynomial solvability of \( (P_k(Z)) \). If \( f \) is quasi-concave on \( D \), either of the extreme points \( x^k(\zeta) \) and \( x^k(\zeta) \) of \( D_k(Z) \) is optimal to \( (P_k(Z)) \). Hence we can also solve such instances of (P) in polynomial time on the average.

**Acknowledgment**

The author is grateful to anonymous referees for their helpful comments and suggestions.

**References**


Minimization of a Saddle Function


Takahito KUNO:
Institute of Information Sciences and Electronics
University of Tsukuba
Tsukuba, Ibaraki 305, Japan
(E-mail: takahito@is.tsukuba.ac.jp)