Global Minimization of a Generalized Convex Multiplicative Function

HIROSHI KONNO[1], TAKAHITO KUNO[2], and YASUTOSHI YAJIMA[3]
[1]*Institute of Human and Social Sciences, Tokyo Institute of Technology*;
[2]*Institute of Information Sciences and Electronics, University of Tsukuba*;
[3]*Department of Industrial Engineering and Management, Tokyo Institute of Technology, Japan*

(Received: 11 July 1991; accepted: 25 May 1993)

Abstract. This paper discusses an algorithm for generalized convex multiplicative programming problems, a special class of nonconvex minimization problems in which the objective function is expressed as a sum of p products of two convex functions. It is shown that this problem can be reduced to a concave minimization problem with only $2p$ variables. An outer approximation algorithm is proposed for solving the resulting problem.

Key words. Nonconvex minimization, global optimization, convex multiplicative function, outer approximation method.

1. Introduction

In this paper, we propose a practical algorithm for solving a *generalized convex multiplicative programming problem*:

$$\left|\begin{array}{ll}\text{minimize} & g(x)+\sum_{i=1}^{p} f_i(x)g_i(x)\\ \text{subject to} & x\in X\,,\end{array}\right. \tag{1.1}$$

where g, f_i's and g_i's are convex functions defined on R^n and $X \subset R^n$ is a compact convex set. This problem has applications in computational geometry [8, 11] and VLSI chip design [14]. Also, as shown in [9], general quadratic programming problems can be put into this form.

The problem (1.1) is a generalization of a *convex multiplicative programming problem*, i.e., a minimization of the product of convex functions over a convex set. The authors studied this type of nonconvex minimization problems in a series of articles [6, 7, 10, 12, 13]. In [6, 12], we treated a special case of (1.1) in which $p = 1$. We proposed a discrete approximation method [6] and a parametric successive underestimation method [12], and demonstrated that both of these methods can solve a fairly large scale problems. Also, the papers [1, 2, 15, 18] deal with this type of problems.

The organization of this paper is as follows: in Section 2, we reduce the problem (1.1) into a $2p$-dimensional concave minimization problem by introducing auxiliary variables, and we look into the structure of the resulting problem. In Section 3, we construct an outer approximation algorithm. This algorithm uses cutting planes exploiting the special structure of the problem. Results of computational experiments of this algorithm are reported in Section 4. In Section 5, we briefly discuss the applications of our method to nonconvex quadratic programming problems and generalized linear fractional programming problems.

2. Reduction of the Problem into a 2*p*-Dimensional Concave Program

Let us consider the generalized convex multiplicative programming problem:

$$(\mathrm{P})\ \left|\begin{array}{ll} \text{minimize} & f(x) = g(x) + \sum_{i=1}^{p} f_i(x) g_i(x) \\ \text{subject to} & x \in X , \end{array}\right. \tag{2.1}$$

where $g,\ f_i,\ g_i: R^n \to R^1$ $(i = 1, \dots, p)$ are convex functions and $X \subset R^n$ is a nonempty, compact and convex set. The product of two convex functions is not convex in general [7, 12], so that the objective function f need not be (quasi)convex. We assume in the sequel that

$$f_i(x) > 0 , \quad g_i(x) > 0 \quad \forall x \in X , \quad i = 1, \dots, p . \tag{2.2}$$

Let us note that, if f_i and g_i are affine, then (2.2) can be assumed without loss of generality. To see this, let

$$v_i < \min\{\min\{f_i(x) \mid x \in X\}, \min\{g_i(x) \mid x \in X\}\} , \quad i = 1, \dots, p , \tag{2.3}$$

and define $\tilde{f}_i(x) = f_i(x) - v_i$, $\tilde{g}_i(x) = g_i(x) - v_i$ and $\tilde{g}(x) = g(x) + \Sigma_{i=1}^{p} [v_i f_i(x) + v_i g_i(x) - v_i^2]$. Then,

$$f(x) = \tilde{g}(x) + \sum_{i=1}^{p} \tilde{f}_i(x) \tilde{g}_i(x) , \tag{2.4}$$

where $\tilde{f}_i$ and $\tilde{g}_i$ satisfy (2.2) and $\tilde{g}$ is still convex.

Let us introduce $2p$ auxiliary variables ζ_i, η_i $(i = 1, \dots, p)$ and define the following problem:

$$\left|\begin{array}{ll} \text{minimize} & F(x, \zeta, \eta) = g(x) + \frac{1}{2} \sum_{i=1}^{p} [\zeta_i (f_i(x))^2 + \eta_i (g_i(x))^2] \\ \text{subject to} & x \in X , \\ & \zeta_i \eta_i \geq 1, \quad i = 1, \dots, p , \\ & (\zeta, \eta) \geq 0 , \end{array}\right. \tag{2.5}$$

where $\zeta = (\zeta_1, \dots, \zeta_p)^t$, $\eta = (\eta_1, \dots, \eta_p)^t$. The objective function F is continuous and bounded from below on the feasible region. Hence (2.5) has a finite optimal solution.

THEOREM 2.1. *Let* (x^*, ζ^*, η^*) *be an optimal solution of* (2.5). *Then* x^* *is an optimal solution of* (P) *and* $f(x^*) = F(x^*, \zeta^*, \eta^*)$.

Proof. For an arbitrary $x \in X$ we have

$$\begin{aligned}
&\min\{F(x, \zeta, \eta) \mid \zeta_i\eta_i \geq 1, i = 1, \dots, p, (\zeta, \eta) \geq 0\} \\
&= g(x) + \frac{1}{2}\sum_{i=1}^{p} \min\{\zeta_i(f_i(x))^2 + \frac{1}{\zeta_i}(g_i(x))^2 \mid \zeta_i > 0\} \\
&= g(x) + \sum_{i=1}^{p} f_i(x)g_i(x)
\end{aligned}$$

by noting (2.2). □

This transformation (2.5) is an extension of the one proposed in [6] for a special case of (P), in which $p = 1$.

For any fixed $(\zeta, \eta) \geq 0$, let us consider a subproblem of (2.5):

$$(\mathrm{P}(\zeta, \eta)) \left| \begin{array}{ll} \text{minimize} & F(x; \zeta, \eta) = g(x) + \frac{1}{2}\sum_{i=1}^{p} [\zeta_i(f_i(x))^2 + \eta_i(g_i(x))^2] \\ \text{subject to} & x \in X\,. \end{array}\right. \tag{2.6}$$

LEMMA 2.2. $F(\cdot; \zeta, \eta)$ *is a convex function for any* $(\zeta, \eta) \geq 0$.

Proof. Follows from Theorem 5.1 in [17]. □

We can obtain an optimal solution $x^*(\zeta, \eta)$ of $(\mathrm{P}(\zeta, \eta))$ by using any one of standard convex minimization algorithms. Let

$$G(\zeta, \eta) = F(x^*(\zeta, \eta), \zeta, \eta)\,. \tag{2.7}$$

Then (2.5) is reduced to a problem of the $2p$ variables (ζ, η):

$$\left| \begin{array}{ll} \text{minimize} & G(\zeta, \eta) \\ \text{subject to} & \zeta_i\eta_i \geq 1, i = 1, \dots, p\,, \\ & (\zeta, \eta) \geq 0\,. \end{array}\right. \tag{2.8}$$

THEOREM 2.3. *G is a concave function and satisfies*

$$G(\zeta^1, \eta^1) \leq G(\zeta^2, \eta^2) \quad \text{if} \quad (\zeta^1, \eta^1) \leq (\zeta^2, \eta^2)\,. \tag{2.9}$$

Proof. Since $F(x, \cdot, \cdot)$ is affine for any fixed x, the function G is the pointwise minimum of a family of affine functions. This implies that G is a concave function. The relation (2.9) is obvious from the definition. □

3. An Outer Approximation Method for the Master Problem

Let us proceed to the algorithm for obtaining a globally optimal solution (ζ^*, η^*) of the concave minimization problem (2.8).

We have the following lemma under the assumption (2.2):

LEMMA 3.1. *The set of optimal solutions of* (2.8) *is bounded.*

Proof. Let (ζ^*, η^*) be an optimal solution of (2.8). It is easy to see that there exists $x^* \in X$ such that $\zeta_i^* = g_i(x^*)/f_i(x^*)$ and $\eta_i^* = f_i(x^*)/g_i(x^*)$. Therefore, both ζ_i^* and η_i^* are bounded, because f_i and g_i defined on R^n are bounded and positive valued on the compact set X. □

For $i = 1, \dots, p$ let $\underline{\zeta}_i$ *and* $\underline{\eta}_i$ be lower bounds of ζ_i^* and η_i^*, respectively. Also let

$$\Psi = \{(\zeta, \eta) \in R^p \times R^p \mid \zeta_i \eta_i \geq 1, i = 1, \dots, p, (\zeta, \eta) \geq 0\}, \tag{3.1}$$

$$\Omega_0 = \{(\zeta, \eta) \in R^p \times R^p \mid \underline{\zeta} \leq \zeta \leq \bar{\zeta}, \underline{\eta} \leq \eta \leq \bar{\eta}\}, \tag{3.2}$$

where

$$\begin{cases} \underline{\zeta} = (\underline{\zeta}_1, \dots, \underline{\zeta}_p)^t, & \bar{\zeta} = (1/\underline{\eta}_1, \dots, 1/\underline{\eta}_p)^t, \\ \underline{\eta} = (\underline{\eta}_1, \dots, \underline{\eta}_p)^t & \bar{\eta} = (1/\underline{\zeta}_1, \dots, 1/\underline{\zeta}_p)^t. \end{cases} \tag{3.3}$$

Then (2.8) is equivalent to the following:

$$(\text{MP}) \left| \begin{array}{ll} \text{minimize} & G(\zeta, \eta) \\ \text{subject to} & (\zeta, \eta) \in \Psi \cap \Omega_0 . \end{array} \right. \tag{3.4}$$

The feasible region $\Psi \cap \Omega_0$ is nonempty, convex and compact. Therefore, we can apply an outer approximation method to (MP) with the initial relaxed problem:

$$(\text{P}_0) \left| \begin{array}{ll} \text{minimize} & G(\zeta, \eta) \\ \text{subject to} & (\zeta, \eta) \in \Omega_0 . \end{array} \right. \tag{3.5}$$

Note that an optimal solution (ζ^0, η^0) of (P_0) is given by $\zeta^0 = \underline{\zeta}$, $\eta^0 = \underline{\eta}$, since G is nondecreasing in each argument (Theorem 2.3). We need to solve a sequence of relaxed problems:

$$(\text{P}_k) \left| \begin{array}{ll} \text{minimize} & G(\zeta, \eta) \\ \text{subject to} & (\zeta, \eta) \in \Omega_k , \end{array} \right. \quad k = 1, 2, \dots, \tag{3.6}$$

such that $\Omega_0 \supset \Omega_1 \supset \Omega_2 \supset \cdots \supset \Psi \cap \Omega_0$. The kth approximation Ω_k of $\Psi \cap \Omega_0$ is generated by adding some constraint $l_{k-1}(\zeta, \eta) \geq 0$ to the system defining Ω_{k-1}, i.e.,

$$\Omega_k = \Omega_{k-1} \cap \{(\zeta, \eta) \in R^p \times R^p \mid l_{k-1}(\zeta, \eta) \leq 0\}\,, \quad k = 1, 2, \ldots . \tag{3.7}$$

If an optimal solution (ζ^k, η^k) of (P_k) is a point of Ψ, then it is a globally optimal solution of (MP), and an optimal solution $x^*(\zeta^k, \eta^k)$ of $(\mathrm{P}(\zeta^k, \eta^k))$ solves the original problem (P).

3.1. CUTTING FUNCTION

For $i = 1, \ldots, p$ let

$$\Psi_i = \{(\zeta_i, \eta_i) \mid \zeta_i \eta_i \geq 1, (\zeta_i, \eta_i) \geq 0\}\,, \tag{3.8}$$

$$\Omega_{i0} = \{(\zeta_i, \eta_i) \mid \underline{\zeta}_i \leq \zeta_i \leq \bar{\zeta}_i\}\,. \tag{3.9}$$

Then the feasible region $\Psi \cap \Omega_0$ of (MP) can be decomposed into the following form:

$$\Psi \cap \Omega_0 = (\Psi_1 \cap \Omega_{10}) \times (\Psi_2 \cap \Omega_{20}) \times \cdots \times (\Psi_p \cap \Omega_{p0})\,. \tag{3.10}$$

To approximate $\Psi \cap \Omega_0$, we may approximate each $\Psi_i \cap \Omega_{i0}$ in the ζ_i-η_i space.

Let (ξ^k, η^k) be an optimal solution of the kth relaxed problem (P_k) and let

$$(\zeta_t^k, \eta_t^l) \in \operatorname{argmin}\{\zeta_i \eta_i \mid (\zeta_i, \eta_i) = (\zeta_1^k, \eta_1^k), \ldots, (\zeta_p^k, \eta_p^k)\}\,. \tag{3.11}$$

We define the cutting function l_k as follows:

$$l_k(\zeta, \eta) = 2 - \zeta_t \sqrt{\eta_t^k / \zeta_t^k} - \eta_t \sqrt{\zeta_t^k / \eta_t^k}\,. \tag{3.12}$$

LEMMA 3.2. *If* $(\zeta^k, \eta^k) \not\in \Psi$, *then*

$$l_k(\zeta, \eta) \leq 0 \quad \forall (\zeta, \eta) \in \Psi \quad \textit{and} \quad l_k(\zeta^k, \eta^k) > 0\,. \tag{3.13}$$

Proof. If $(\zeta, \eta) \in \Psi$, then $\zeta_i \eta_i \geq 1$ for each i, and hence

$$l_k(\zeta, \eta) \leq 2 - 2\sqrt{\zeta_t \sqrt{\eta_t^k / \zeta_t^k} \cdot \eta_t \sqrt{\zeta_t^k / \eta_t^k}} \leq 0\,.$$

Also we have

$$l_k(\zeta^k, \eta^k) = 2 - 2\sqrt{\zeta_t^k \eta_t^k} > 0\,,$$

by noting $\zeta_t^k \eta_t^k < 1$. □

In the ζ_t-η_t space, the set $\{(\zeta_t, \eta_t) \mid l_k(\zeta, \eta) = 0\}$ is a supporting hyperplane of Ψ_t at $(\sqrt{\zeta_t^k / \eta_t^k}, \sqrt{\eta_t^k / \zeta_t^k})$, which is the intersection of the boundary of Ψ_t and the ray emanating from the origin to the point (ζ_t^k, η_t^k) (see Figure 2 in Section 3.4).

The normal of l_k is orthogonal to every ζ_i-η_i space except the ζ_t-η_t space. If we define Ω_{k+1} according to (3.7), it must be expressed as

$$\Omega_{k+1} = \Omega_{1,k+1} \times \Omega_{2,k+1} \times \cdots \times \Omega_{p,k+1}\,, \quad k = 0, 1, 2 \ldots, \tag{3.14}$$

where $\Omega_{i,k+1} \subset R^2$ is a polytope and satisfies

$$\Omega_{tk} \supset \Omega_{t,k+1} \supset \Psi_t \cap \Omega_{t0} \quad \text{and} \quad \Omega_{ik} = \Omega_{i,k+1} \supset \Psi_i \cup \Omega_{i0}\,, \quad i \neq t\,. \tag{3.15}$$

In the whole space, the ray emanating from the origin to the point (ζ^k, η^k) always intersects $\Psi \cap \Omega_0$. The cutting function l_k defined by (3.12) can also be derived from this property in the framework of the ordinary outer approximation method [3, 4].

3.2. ALGORITHM

We are now ready to construct an outer approximation algorithm for solving (MP). Let $\epsilon \geq 0$ be a give tolerance.

ALGORITHM OAM

Step 0. Let $k = 0$.

Step 1. Compute an optimal solution (ζ^k, η^k) of (P_k) and let $\zeta_i^\epsilon = \sqrt{\zeta_i^k/\eta_i^k}$, $\eta_i^\epsilon = \sqrt{\eta_i^k/\zeta_i^k}$, $i = 1, \ldots, p$.

Step 2. Let $(\zeta_t^k, \eta_t^k) \in \operatorname{argmin}\{\zeta_i\eta_i \mid (\zeta_i, \eta_i) = (\zeta_1^k, \eta_1^k), \ldots, (\zeta_p^k, \eta_p^k)\}$. If

$$1 - \zeta_t^k \eta_t^k \leq \epsilon\,, \tag{3.16}$$

then stop.

Step 3. Let $l_k(\zeta, \eta) = 2 - \eta_t^\epsilon \zeta_t - \zeta_t^\epsilon \eta_t$. Update the feasible region as $\Omega_{k+1} = \Omega_k \cap \{(\zeta, \eta) \in R^p \times R^p \mid l_k(\zeta, \eta) \leq 0\}$.

Step 4. Let $k = k + 1$ and return to Step 1. □

THEOREM 3.3 *If $\epsilon > 0$, then Algorithm OAM terminates after finitely many iterations and yields an approximate solutions $(\zeta^\epsilon, \eta^\epsilon)$. If $\epsilon = 0$, then OAM generates a sequence $\{(\zeta^k, \eta^k)\}$, every accumulation point of which is a globally optimal solution of (MP).*

Proof. Assume that Algorithm OAM is infinite. Then there exists a subsequence $\{(\zeta^{k_q}, \eta^{k_q})\}$ such that

$$1 - \zeta_t^{k_q} \eta_t^{k_q} > \epsilon \quad \forall q\,, \tag{3.17}$$

where the index t is determined by (3.11). Since all (ζ^k, η^k)'s are generated in the compact set Ω_0, we may assume that $\{(\zeta^{k_q}, \eta^{k_q})\}$ converges to some point $(\tilde{\zeta}, \tilde{\eta})$. Let

$$\tilde{l}(\zeta,\eta)=2-\zeta_t\sqrt{\tilde{\eta}_t/\tilde{\zeta}_t}-\eta_t\sqrt{\tilde{\zeta}_t/\tilde{\eta}_t}\,. \tag{3.18}$$

For every q, we have $(\zeta^{k_{q+1}},\eta^{k_{q+1}})\in\Omega_{k_{q+1}}\subset\Omega_{k_q}$ and hence $l_{k_q}(\zeta^{k_{q+1}},\eta^{k_{q+1}})<0$. Thus,

$$\lim_{q\to\infty} l_{k_q}(\zeta^{k_{q+1}},\eta^{k_{q+1}})=\lim_{q\to\infty} l_{k_q}(\zeta^{k_q},\eta^{k_q})=\tilde{l}(\tilde{\zeta},\tilde{\eta})\leqslant 0\,.$$

However, by (3.18) we have

$$\tilde{l}(\tilde{\zeta},\tilde{\eta})=2(1-\sqrt{\tilde{\zeta}_t\tilde{\eta}_t})\leqslant 0\,,$$

which contradicts (3.17). If $\epsilon>0$, therefore, OAM must terminate after finitely many iterations. If $\epsilon=0$, then

$$G(\tilde{\zeta},\tilde{\eta})=\lim_{q\to+\infty} G(\zeta^{k_q},\eta^{k_q})\leqslant G(\zeta^*,\eta^*)\,,$$

because $G(\zeta^{k_q},\eta^{k_q})\leqslant G(\zeta^*,\eta^*)$ for every q. □

An approximate solution $x^*(\zeta^\epsilon,\eta^\epsilon)$ of the original problem (P) can be obtained by solving $(\mathrm{P}(\zeta^\epsilon,\eta^\epsilon))$. If the stopping criterion (3.16) of Algorithm OAM is replaced by the following:

$$G(\zeta^\epsilon,\eta^\epsilon)-G(\zeta^k,\eta^k)\leqslant\epsilon\,, \tag{3.19}$$

then we will obtain a globally ϵ-optimal solution of (MP) and that of (P). If $\epsilon>0$, we can prove the finiteness of the algorithm analogously as above.

3.3. SOLUTION OF THE RELAXED PROBLEM

We have to solve a relaxed problem (P_k) in each iteration of Algorithm OAM. Since (P_k) is a concave minimization problem, a globally optimal solution (ζ^k,η^k) exists among the vertices $V(\Omega_k)$ of its feasible region Ω_k. Therefore we can find (ζ^k,η^k) by solving $(\mathrm{P}(\zeta,\eta))$ for every $(\zeta,\eta)\in V(\Omega_k)$. Let V_k be the set of vertices newly generated by adding the constraint $l_k(\zeta,\eta)\leqslant 0$ to the system defining Ω_k. Then we have

$$V(\Omega_{k+1})=V_k\cup\{(\zeta,\eta)\in V(\Omega_k)\mid l_k(\zeta,\eta)\leqslant 0\}\,. \tag{3.20}$$

The efficiency of the algorithm depends strongly upon the computation of V_k.

Recall that the feasible region Ω_k of (P_k) is the orthogonal product of polytopes Ω_{ik}'s defined in their respective ζ_i-η_i spaces. Hence the vertices of Ω_k can also be expressed as follows:

$$V(\Omega_k)=V(\Omega_{1k})\times V(\Omega_{2k})\times\cdots\times V(\Omega_{pk})\,. \tag{3.21}$$

PROPOSITION 3.4. $\{(\zeta_t, \eta_t) \mid l_k(\zeta, \eta) = 0\}$ *supports* $\Psi_t \cap \Omega_{t0}$.

Proof. The hyperplane $\{(\zeta_t, \eta_t) \mid l_k(\zeta, \eta) = 0\}$ supports Ψ_t and $(\sqrt{\zeta_t^k/\eta_t^k}, \sqrt{\eta_t^k/\zeta_t^k})$. Since $(\zeta^k, \eta^k) \in \Omega_0$, it follows from (3.2) and (3.3) that

$$\underline{\zeta}_t \leq \zeta_t^k \leq 1/\underline{\eta}_t, \quad \underline{\eta}_t \leq \eta_t^k \leq 1/\underline{\zeta}_t .$$

Thus, $(\sqrt{\zeta_t^k/\eta_t^k}, \sqrt{\eta_t^k/\zeta_t^k})$ is contained in Ω_{t0}, and hence $\{(\zeta_t, \eta_t) \mid l_k(\zeta, \eta) = 0\}$ supports $\Psi_t \cap \Omega_{t0}$. □

This proposition guarantees that redundant constraints cannot occur in each ζ_i-η_i space. Therefore, l_k cuts off exactly one vertex (ζ_t^k, η^k) from Ω_{tk} and generates $\Omega_{t,k+1}$ with two new vertices, say (ζ_t', η_t') and (ζ_t'', η_t''). On the other hand, we have $\Omega_{i,k+1} = \Omega_{ik}$ for every $i \neq t$. Consequently, we have

$$\begin{aligned} V_k = {} & V(\Omega_{1k}) \times \cdots \times V(\Omega_{t-1,k}) \\ & \times \{(\zeta_t', \eta_t'), (\zeta_t'', \eta_t'')\} \times V(\Omega_{t+1,k}) \times \cdots \times V(\Omega_{pk}) . \end{aligned} \tag{3.22}$$

Although $|V_k|$ might be a large number, we can compute V_k without any expensive procedures.

3.4. NUMERICAL EXAMPLE

Before concluding this section, let us illustrate Algorithm OAM by using the following two-dimensional problem:

$$\left|\begin{array}{ll} \text{minimize} & f(x) = 3x_1 - 4x_2 + (x_1 + 2x_2 - 1.5)(2x_1 - x_2 + 4) \\ & \qquad + (x_1 - 2x_2 + 8.5)(2x_1 + x_2 - 1) \\ \text{subject to} & 5x_1 - 8x_2 \geq -24, 5x_1 + 8x_2 \leq 44, 6x_1 - 3x_2 \leq 15 , \\ & 4x_1 + 5x_2 \geq 10 , \quad x_1 \geq 0 . \end{array}\right. \tag{3.23}$$

We see from Figure 1 that

$$\begin{cases} 1 \leq x_1 + 2x_2 - 1.5 \leq 9 , & 1 \leq 2x_1 - x_2 + 4 \leq 9 , \\ 2 \leq x_1 - 2x_2 + 8.5 \leq 11 , & 1 \leq 2x_1 + x_2 - 1 \leq 10 , \end{cases} \tag{3.24}$$

for all x in the feasible region X. Thus the assumption (2.2) is satisfied. The objective function value $G(\zeta, \eta)$ of (MP) associated with (3.23) is given by solving a convex quadratic program:

$$\left|\begin{array}{ll} \text{minimize} & F(x; \zeta, \eta) = 3x_1 - 4x_2 + \frac{1}{2}[\zeta_1(x_1 + 2x_2 - 1.5)^2 \\ & \qquad + \eta_1(2x_1 - x_2 + 4)^2 + \zeta_2(x_1 - 2x_2 + 8.5)^2 \\ & \qquad + \eta_2(2x_1 + x_2 - 1)^2] \end{array}\right. \tag{3.25}$$

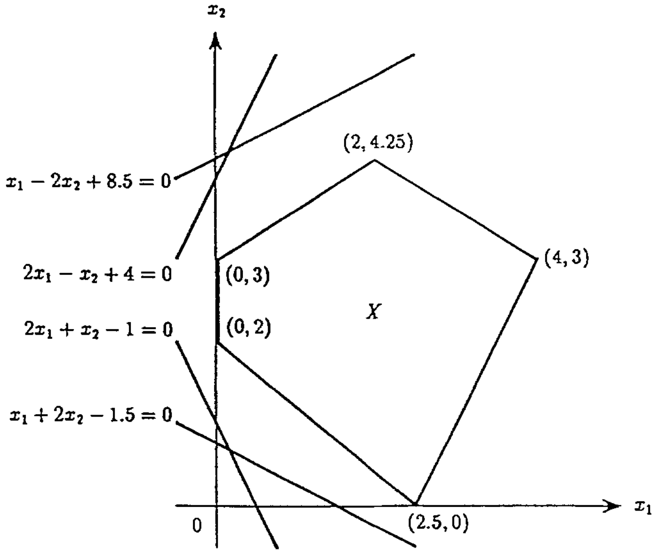

Fig. 1. Example (3.23) described in Section 3.4.

subject to $x \in X$.

By (3.24) we define the bounds of ζ_i^*'s and η_i^*'s below:

$$\underline{\zeta}_1 = \tfrac{1}{9} = 0.111, \quad \overline{\zeta}_1 = \tfrac{9}{1} = 9.000, \quad \underline{\eta}_1 = \tfrac{1}{9} = 0.111, \quad \overline{\eta}_1 = \tfrac{9}{1} = 9.000,$$
$$\underline{\zeta}_2 = \tfrac{1}{11} = 0.091, \quad \overline{\zeta}_2 = \tfrac{10}{2} = 5.000, \quad \underline{\eta}_2 = \tfrac{2}{10} = 0.200, \quad \overline{\eta}_2 = \tfrac{11}{1} = 11.000.$$

Then the feasible region of the initial relaxed problem (P_0) is as follows:

$$\Omega_0 = \{(\zeta_1, \eta_1) \mid 0.1111 \leq \zeta_1 \leq 9.000, 0.111 \leq \eta_1 \leq 9.000\}$$
$$\times \{(\zeta_2, \eta_2) \mid 0.091 \leq \zeta_2 \leq 5.000, 0.200 \leq \eta_2 \leq 11.000\}$$

(see Figure 2). The function G attains its minimum (ζ^0, η^0) over Ω_0 at

$$(\zeta_1, \zeta_2, \eta_1, \eta_2) = (0.111, 0.091, 0.111, 0.200).$$

The value of G at this point is

$$\min\{F(x; 0.111, 0.091, 0.111, 0.200) \mid x \in X\} = -10.135$$

(see Figure 3). Since $0.111 \cdot 0.111 = 0.012 < 0.091 \cdot 0.200 = 0.018$, we define

$$l_0(\zeta, \eta) = 2 - \sqrt{0.111/0.111}\zeta_1 - \sqrt{0.111/0.111}\eta_1 = 2 - \zeta_1 - \eta_1.$$

Then we obtain the first approximation of $\Psi \cap \Omega_0$:

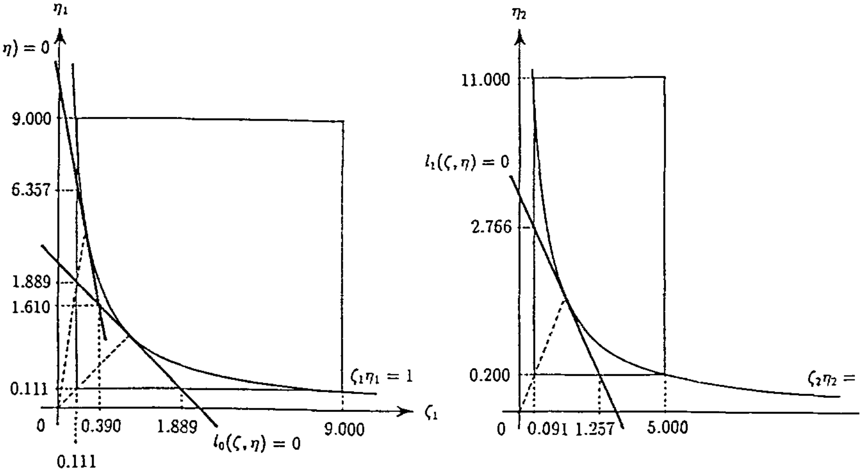

Fig. 2. Relation between the cut l_k and the set Ω_k.

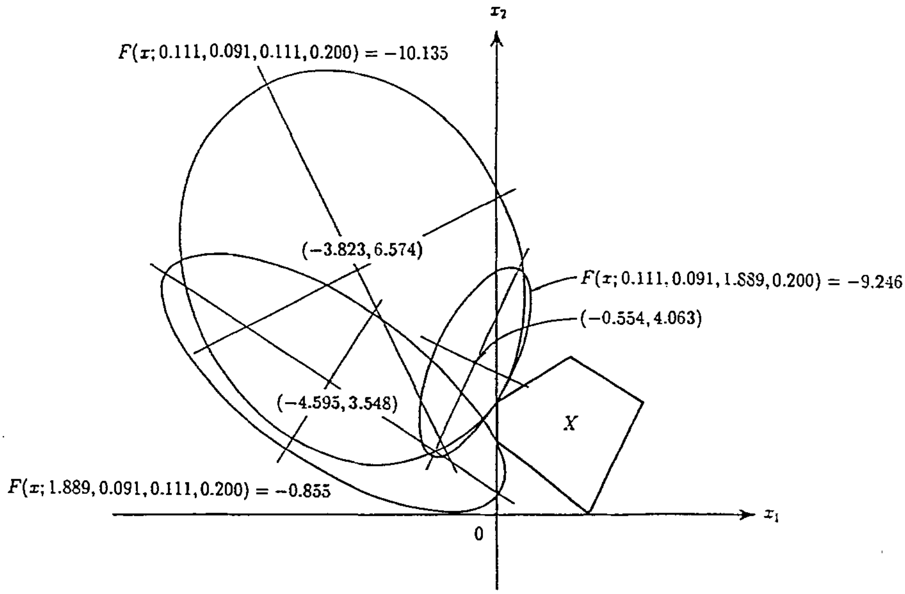

Fig. 3. Calculating $G(\zeta, \eta)$ for the example (3.23).

$$\begin{aligned}\Omega_1 = \{(\zeta_1, \eta_1) \mid \zeta_1 + \eta_1 \geqslant 2, 0.111 \leqslant \zeta_1 \leqslant 9.000, 0.111 \leqslant \eta_1 \leqslant 9.000\} \\ \times \{(\zeta_2, \eta_2) \mid 0.091 \leqslant \zeta_2 \leqslant 5.000, 0.200 \leqslant \eta_2 \leqslant 11.000\}\,.\end{aligned}$$

The constraint $\zeta_1 + \eta_1 \geqslant 2$ cuts the vertex (0.111, 0.091, 0.111, 0.200) from Ω_0 and generates two new vertices of Ω_1 (Figure 2):

$$\begin{aligned}&(\zeta_1, \zeta_2, \eta_1, \eta_2) \\ &= (1.889, 0.091, 0.111, 0.200), (0.111, 0.091, 1.889, 0.200)\,.\end{aligned}$$

The values of G at these points are

$$\min\{F(x; 1.889, 0.091, 0.111, 0.200) \mid x \in X\} = -0.855\,,$$
$$\min\{F(x; 0.111, 0.091, 1.889, 0.200) \mid x \in X\} = -9.246\,,$$

respectively (Figure 3). Thus, an optimal solution (ζ^1, η^1) of the relaxed problem (P_1) is (0.111, 0.091, 1.889, 0.200). Since $0.111 \cdot 1.889 = 0.210 > 0.091 \cdot 0.200 = 0.018$, we define

$$l_1(\zeta, \eta) = 2 - \sqrt{0.200/0.091}\zeta_2 - \sqrt{0.091/0.200}\eta_2 = 2 - 1.483\zeta_2 - 0.674\eta_2\,,$$

and

$$\begin{aligned}\Omega_2 = \{(\zeta_1, \eta_1) \mid \zeta_1 + \eta_1 \geq 2, 0.111 \leq \zeta_1 \leq 9.000, 0.111 \leq \eta_1 \leq 9.000\}\\ \times \{(\zeta_2, \eta_2) \mid 1.483\zeta_2 + 0.674\zeta_2 \geq 2, 0.091 \leq \zeta_2 \leq 5.000,\\ 0.200 \leq \eta_2 \leq 11.000\}\,.\end{aligned}$$

Two vertices (1.889, 0.091, 0.111, 0.200) and (0.111, 0.091, 1.889, 0.200) are cut off and the following vertices are newly generated (Figure 2):

$$\begin{aligned}(\zeta_1, \chi_2, \eta_1, \eta_2) = (1.889, 1.257, 0.111, 0.200), (0.111, 1.257, 1.889, 0.200)\,,\\ (1.889, 0.091, 0.111, 2.766), (0.111, 0.091, 1.889, 2.766)\,.\end{aligned}$$

An optimal solution (ζ^2, η^2) of the relaxed problem (P_2) is (0.111, 1.257, 1.889, 0.200) and its optimal value is -5.601. Since $0.111 \cdot 1.889 = 0.210 < 1.257 \cdot 0.200 = 0.252$, we define

$$l_2(\zeta, \eta) = 2 - \sqrt{1.889/0.111}\zeta_1 - \sqrt{0.111/1.889}\eta_1 = 2 - 4.123\zeta_1 - 0.243\eta_1\,.$$

We obtain Ω_3 by adding the constraint $4.123\zeta_1 + 0.243\eta_1 \geq 2$ to the system defining Ω_2.

Two vertices (0.111, 1.257, 1.889, 0.200) and (0.111, 0.091, 1.889, 2.766) are cut off and four new vertices are generated (Figure 2):

$$\begin{aligned}(\zeta_1, \zeta_2, \eta_1, \eta_2) = (0.390, 1.257, 1.610, 0.200), (0.390, 0.091, 1.610, 2.766)\,,\\ (0.111, 1.257, 6.357, 0.200), (0.111, 0.091, 6.357, 2.766)\,.\end{aligned}$$

An optimal solution (ζ^3, η^3) of (P_3) is (0.111, 1.257, 6.357, 0.200) and its optimal value is -3.367.

In this way, a sequence $\{(\zeta^k, \xi^k)\}$ will be generated. Its accumulation point $(\zeta^*, \eta^*) = (0.222, 0.800, 4.500, 1.250)$ is a globally optimal solution of (MP). We obtain an optimal solution $x^* = (0.000, 3.000)$ of (3.23) and its optimal value -2.5 by solving (3.25) with $\zeta = \zeta^*$, $\eta = \eta^*$.

4. Computational Experiments

We report the results of computational experiments on the algorithm presented in the previous section. We solved two subclasses of the forms:

$$(\text{TP1}) \left| \begin{array}{ll} \text{minimize} & c_0^t x + \frac{1}{2} x^t Q^t Q x + \sum_{i=1}^{p} c_i^t x d_i^t x \\ \text{subject to} & Ax \geqslant b\,, \quad x \geqslant 0\,, \end{array} \right.$$

$$(\text{TP2}) \left| \begin{array}{ll} \text{minimize} & \sum_{i=1}^{p} c_i^t x d_i^t x \\ \text{subject to} & Ax \geqslant b\,, \quad x \geqslant 0\,, \end{array} \right.$$

where c_0, c_i, $d_i \in R^n$ $(i = 1, \ldots, p)$, $Q \in R^{n \times n}$, $A \in R^{m \times n}$ and $b \in R^m$. All elements of c_i's, d_i's, Q, A and b are randomly generated, whose ranges are $[0, 100]$.

We solved every subproblem ($\mathrm{P}(\zeta, \eta)$) by applying the reduced gradient method [19]. Direction vectors were generated by the conjugate gradient procedure [5]. The size of tolerance ϵ was always fixed at 10^{-5} and both the lower bounds $\underline{\zeta}_i$'s and $\underline{\eta}_i$'s were 10^{-5}. The algorithm was coded in C language and tested on a SUN SPARC-2 computer (27.5 MIPS).

Table I shows the comparison of three algorithms for (TP1) when $p = 1$. Here OAM represents the algorithm presented in Section 3, and PSUM and DAM are the parametric successive underestimation method [12] and the discrete approximation method [6], respectively. For each size of (m, n), the table contains the average CPU time in seconds and the average number of cuts (and their respective standard deviations in the brackets) needed for solving ten examples. Also the average number of vertices generated by cuts in the course of computation (and its standard deviation) is listed in it. This number corresponds to that of subproblems solved for one example. Both the results of PSUM and

Table I. Results of three algorithms for (TP1) when $p = 1$

m	10	30	30	70	70	130	130
n	20	20	50	50	100	100	150
Average CPU time in seconds (standard deviation)							
OAM:	0.5	1.9	8.7	27.9	83.5	288.2	482.2
	(0.1)	(1.0)	(4.0)	(13.6)	(28.7)	(99.9)	(131.7)
PSUM:	1.86	7.69	40.12	174.83	614.62	1002.81	–
DAM:	0.99	3.26	16.42	55.46	229.65	511.42	–
Average # of cuts (standard deviation)							
OAM:	8.9	10.6	10.5	10.2	8.1	11.7	11.5
	(3.5)	(3.0)	(2.7)	(4.3)	(3.4)	(4.2)	(3.5)
Average # of vertices (standard deviation)							
OAM:	15.8	19.2	19.0	18.4	14.2	21.4	21.0
	(7.0)	(6.0)	(5.3)	(8.7)	(6.8)	(8.5)	(7.1)

Table II. Results of OAM for (TP1)

p	2	2	2	3	3	3
m	30	30	70	30	30	70
n	20	50	50	20	50	50
Average CPU time in seconds (standard deviation)						
	7.0	56.2	183.8	421.4	536.0	2101.4
	(4.1)	(43.0)	(57.4)	(1096.5)	(193.1)	(881.4)
Average # of cuts (standard deviation)						
	19.6	22.2	23.9	28.1	33.0	35.1
	(3.0)	(3.8)	(3.0)	(2.9)	(4.1)	(5.2)
Average # of vertices (standard deviation)						
	212.8	270.8	308.8	2048.0	3193.0	3896.6
	(65.2)	(96.9)	(80.0)	(625.0)	(1089.4)	(1816.5)

DAM are taken from [12], in which their experiments were carried out on a SUN 4/280S computer (8.5 MIPS). Tables II and III show the results of OAM for (TP1) and (TP2), respectively, when (p, m, n) ranges from $(2, 30, 20)$ to $(4, 70, 50)$. The average CPU time and the average numbers of cuts and vertices of ten examples for each (p, m, n) are listed in them.

We see from these tables that Algorithm OAM is very sensitive to the size of p. Both the numbers of cuts and vertices generated through computation sharply increase as functions of p. As expected from (3.22), the latter is rather conspicuous for this tendency compared with the former. However, it should be emphasized that these numbers slowly increase for each p as the size of (m, n) gets larger.

When p is fixed at a small number, say $p \leq 3$, OAM is reasonable efficient. In particular when $p = 1$, OAM solves (TP1) in about half computational time required by the parametric successive underestimation method (PSUM), even after taking the difference of their experimental environments into consideration. In this case, the total computational time is dominated by that needed for solving the associated convex quadratic program, i.e., $(\mathrm{P}(\zeta, \eta))$. We have to devise more efficient algorithm for convex programs when the size of (m, n) is larger.

Table III. Results of OAM for (TP2)

p	2	2	2	3	3	3	4	4
m	30	30	70	30	30	70	30	30
n	20	50	50	20	50	50	20	50
Average CPU time in seconds (standard deviation)								
	5.4	25.9	55.6	49.3	202.7	1087.7	416.5	3897.6
	(1.8)	(5.1)	(14.8)	(33.1)	(74.2)	(900.4)	(233.2)	(2158.6)
Average # of cuts (standard deviation)								
	21.2	21.6	19.6	29.5	30.2	32.3	38.8	42.7
	(2.5)	(1.8)	(1.7)	(4.0)	(3.2)	(5.2)	(5.0)	(5.6)
Average # of vertices (standard deviation)								
	246.2	253.0	206.6	2388.4	2509.8	3145.8	25088.8	36682.2
	(57.1)	(41.1)	(31.7)	(1039.8)	(770.8)	(1475.8)	(11591.6)	(18355.2)

5. Some Extensions

Let us consider a quadratic programming problem:

$$\text{(QP)}\left|\begin{array}{ll}\text{minimize} & f(x) = c^t x + \frac{1}{2} x^t Q x \\ \text{subject to} & Ax \geqslant b\,, \quad x \geqslant 0\,,\end{array}\right. \tag{5.1}$$

where $c \in R^n$, $Q \in R^{n \times n}$, $A \in R^{m \times n}$ and $b \in R^m$.

LEMMA 5.1. *If the rank of Q is p ($\leqslant n$), then the objective function of (QP) can be expressed by linearly independent sets of vectors, $\{c_1, \ldots, c_p\}$, $\{d_1, \ldots, d_p\} \subset R^n$, as follows*:

$$f(x) = c^t x + \sum_{i=1}^{p} c_i^t x \cdot d_i^t x\,. \tag{5.2}$$

Proof. Follows from Theorem 2.2 of [9]. □

Thus (QP) can be put into the same form as (P). In this case, as shown in Section 2, we can assume (2.2) without loss of generality. Therefore every quadratic program can be solved by Algorithm OAM. Similarly, we can apply OAM to bilinear programming problems:

$$\text{(BLP)}\left|\begin{array}{lll}\text{minimize} & c^t x + d^t y + x^t Q y & \\ \text{subject to} & A_1 x \geqslant b_1\,, & x \geqslant 0\,, \\ & A_2 y \geqslant b_2\,, & y \geqslant 0\,.\end{array}\right. \tag{5.3}$$

Finally, let us consider a generalized linear fractional programming problem:

$$\text{(LFP)}\left|\begin{array}{ll}\text{minimize} & f(x) = g(x) + \sum_{i=1}^{p} \dfrac{c_i^t x + c_{i0}}{d_i^t x + d_{i0}} \\ \text{subject to} & x \in X\,,\end{array}\right. \tag{5.4}$$

where c_i, $d_i \in R^n$, c_{i0}, $d_{i0} \in R^1$, g: $R^n \to R^1$ is convex function. If

$$c_i^t x + c_{i0} > 0\,, \quad d_i^t x + d_{i0} > 0 \quad \forall x \in X\,, \quad i = 1, \ldots, p \tag{5.5}$$

we can transform (LFP) into the following equivalent problem:

$$\left|\begin{array}{ll}\text{minimize} & F(x, \zeta, \eta) = g(x) + \frac{1}{2} \sum_{i=1}^{p} [\zeta_i (c_i^t x + c_{i0})^2 + \eta_i / (d_i^t x + d_{i0})^2] \\ \text{subject to} & x \in X\,, \\ & \zeta_i \geqslant 1\,, \quad i = 1, \ldots, p\,, \\ & \zeta \geqslant 0\,, \quad \eta \geqslant 0\,.\end{array}\right. \tag{5.6}$$

It is easy to check (see Theorem 4.2 of [6]) that a subproblem of (5.6):

$$\left| \begin{array}{ll} \text{minimize} & F(x; \zeta, \eta) = g(x) + \frac{1}{2} \sum_{i=1}^{p} [\zeta_i (c_i^t x + c_{i0})^2 + \eta_i / (d_i^t x + d_{i0})^2] \\ \text{subject to} & x \in X \end{array} \right. \tag{5.7}$$

is a convex minimization problem. Hence OAM can be applied to (LFP) as well.

Acknowledgment

The authors are grateful to Professor N. V. Thoai for his helpful comments and suggestions.

This research was partly supported by Grant-in-Aid for Scientific Research of the Ministry of Education, Science and Culture, Grant No. (C)03832018 and (C)04832010.

References

1. Aneja, Y. P., V. Aggarwal, and K. P. K. Nair (1984), On a class of quadratic programming, *European J. of Operational Research* **18**, 62–70.
2. Bector, C. R. and M. Dahl (1974), Simplex type finite iteration technique and reality for a special type of pseudo-concave quadratic functions, *Cahiers du Centre d'Etudes de Recherche Operationnelle* **16**, 207–222.
3. Horst, R., N. V. Thoai, and H. Tuy (1987), Outer approximation by polyhedral convex sets, *OR Spectrum* **9**, 153–159.
4. Horst, R. and H. Tuy (1990), *Global Optimization: Deterministic Approaches*, Springer Verlag.
5. Fletcher, R. and M. Reeves (1964), Function minimization by conjugate gradients, *Computer Journal* **7**, 149–154.
6. Konno, H. and T. Kuno (1990), Generalized linear multiplicative and fractional programming, *Annals of Operations Research* **25**, 147–162.
7. Konno, H. and T. Kuno (1992), Linear multiplicative programming, *Mathematical Programming* **56**, 51–64.
8. Konno, H. T. Kuno, S. Suzuki, P. T. Thach, and Y. Yajima (1991), Global optimization techniques for a problem in the plane, IHSS 91-35, Institute of Human and Social Sciences, Tokyo Institute of Technology.
9. Konno, H. and Y. Yajima (1990), Solving rank two bilinear programs by parametric simplex algorithms, IHSS 90-17, Institute of Human and Social Sciences, Tokyo Institute of Technology.
10. Konno, H., Y. Yajima, and T. Matsui (1991), Parametric simplex algorithms for solving a special class of nonconvex minimization problems, *J. of Global Optimization* **1**, 65–81.
11. Kuno, T. (1991), Globally determining a minimum-area rectangle enclosing the projection of a higher-dimensional set, ISE-TR-91-95, Institute of Information Sciences and Electronics, University of Tsukuba; to appear in *Operations Research Letters*.
12. Kuno, T. and H. Konno (1991), A parameter successive underestimation method for convex multiplicative programming problems, *J. of Global Optimization* **1**, 267–285.
13. Kuno, T., Y. Yajima, and H. Konno (1991), An outer approximation method for minimizing the product of several convex functions on a convex set, IHSS 90-33, Institute of Human and Social Sciences, Tokyo Institute of Technology; also in *J. of Global Optimization* **3** (1993), 325–335.

14. Maling, K., S. H. Mueller, and W. R. Heller (1992), On finding most optimal rectangular package plans, *Proc. of the 19th Design Automation Conference*, 663–670.
15. Pardalos, P. M. (1988), Polynomial time algorithms for some classes of constrained non-convex quadratic problems, *Optimization* **21**, 843–853.
16. Pardalos, P. M. and J. B. Rosen (1987), *Constrained Global Optimization: Algorithms and Applications*, Springer Verlag, Lecture Notes in Computer Science, Vol. 268.
17. Rockafellar, R. T. (1972), *Convex Analysis*, Princeton University Press.
18. Swarup, K. (1966). Indefinite quadratic programming, *Cahiers du Centre d'Etudes de Recherche Operationnelle* **8**, 217–222.
19. Wolfe, P. (1967), Methods of nonlinear programming, in J. Abadie (ed.), *Nonlinear Programming*, North-Holland.